TOWARDS A BETTER UNDERSTANDING OF EARLY DRUG-INDUCED REGULATORY MECHANISMS OF LIVER TUMORIGENESIS

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von
RAPHAËLLE LUISIER
aus
LEYTRON, WALLIS

BASEL, 2014

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Erik Van Nimwegen

Dr Rémi Terranova

Prof. MD. Gerd Kullak-Ublick

Basel, 17 September 2013

Prof. Dr. Jörg Schibler
Dekan

Acknowledgments

The achievement of this PhD research project would not have been possible without the contribution of many people which I would like to acknowledge here.

First I would like to express my gratitude to my dedicated PhD committee: Gerd Kullak-Ublick, for his great external opinion on the project and nice discussions; Rémi Terranova, who gave me the opportunity to undertake this research project at Novartis and who initiated me into the discipline of epigenetics; Erik Van Nimwegen, who constantly challenged me with a critical eye on my project, taught how to rigorously analyze data leading to exciting results, and who's crucial and much appreciated mentoring and out-of-the-box thinking inspired me and provided me with original ideas for this research.

Many colleagues at Novartis contributed to make my life fun and much easier. First I would like to thank Jonathan Moggs, who initiated me into the discipline of toxicology, and gave me the opportunity to take part to MARCAR consortium meetings; Olivier Grenet for nicely introducing me to many relevant people in Novartis and for nice discussions. From DIS I would also like to to thank the 'french group' composed of Gregory, Jean-Philippe, Tulli, Sarah, Audrey D., Virginie, Magali and Julien. Thanks to Natasha: your success in leading both a great scientific career and a family is rousing! To my previous officemates Harri, who nicely initiated me into the problem of phenobarbital-mediated liver NGC and epigenetics; and Nina Scherbichler: it was a delight to supervise such as smiley, fresh and radiant student. DIS is composed of many great scientists which I was lucky to spent some time with: Pierre M., Federico B., Valérie, François P., Daniel S., Philippe C., Ricco, and Axel.

When I arrived in Novartis, I was largely ignorant of the bioinformatic "world", so I profoundly want to thank Arne and Florian: I now have a better understanding of the spirit and philosophy of bioinformatics nerds. Thanks for teaching me the fundamentals of data analysis, R, for debugging my ... nasty bugs. But of course you were much more than two geeky colleagues, so thanks for the fun, the fun and the fun. To David H., I really missed your ironic ton, you left DIS too soon!

Of my colleagues at Novartis, I feel lucky and privileged to have found great and precious friends: Nicole, you rock and it was a delight to have you as my office neighbor; Manuela, un regard bienveillant sans oublier de précieux conseils sur la gastronomie française!; Michèle, une bouffée d'énergie positive à chaque pause en ta compagnie; Audrey K. du rire au professional mentoring, je ne sais comment te remercier pour tes précieux conseils, ton soutien moral et ta vision de battante et de winner!

I also had the opportunity to meet many great people throughout NIBR, particularly thanks to Michael R. who nicely organised the bioinformatics meeting where I got to know exciting bioinformaticians: Sophie B., Caroline G., Ieuan, Yan A., Gugliemo, Ed O; also the great team of bioinformaticians from FMI: Michael S., Dimo, Lukas and Hans-Rudolf, I was really happy to get to know you and only regret to not have spent more time around FMI; I learnt a lot from you guys!

During this PhD I also had the privilege to collaborate with two great experts in the field of toxicology: Jay Goodman and Michael Schwarz. I deeply thank both of you for the interesting discussions, your help and input on the project and your constant availability.

Thanks to Erik, I had the great chance to share some of the Computational Systems biology lab's life and eventually to get to know awesome people: Piotr, Nick, Frederic, Silvia, Luise, Saeed, Matthias, Olin, Peter, Mikhail, Florian, Chris, Philip and Evgeny, thanks to all of you for the welcome, the support and the fun.

I also would like to take here the opportunity to acknowledge great scientists and teachers who, whilst not having directly contributed to this work, inspired me, and have molded me into the researcher I am today: Graeme Pettet, Jeffrey A. Hubbell, Zee Upton, Ruth Luthi-Carter, Melody Swartz and Emile Dupont.

Je tiens à remercier du fond du coeur mes parents, mes soeurs, mes adorables nièces, Coco&Franz pour l'amour, le constant soutien, et les précieuses attentions dont j'ai bénéficié tout au long de mes études et qui m'ont donné la force et l'énergie de continuer. En Valais, j'ai également partagé des moments de rire et de chaleur qui m'ont profondément aidée à maintenir le cap et à relativiser: merci à ma belle-famille, à mes pécieuses amies et à toute la clique.

Je dédie enfin ce travail à mon geeky mentor Tim de Tim avec qui je partage depuis 13 ans les bonnes et les moins bonnes frasques de la vie, qui m'a patiemment écoutée et soutenue dans les moments difficiles, et qui a su trouvé les mots pour me permettre de continuer et surtout aboutir ce travail. Qui plus est, sans Tim, nul doute que je n'aurais jamais goûté au plaisir de la 'compile' sous toutes ses formes, c'est dire si je lui dois tout, ou presque. Pour tout ceci et bien plus encore, du fond du coeur, merci.

Summary

This thesis summarizes the main findings of the research project lead from September 2010 until August 2013 in the laboratory of *Safety Epigenetics* in the Pre-Clinical Safety group (PCS) of the Novartis Institutes for Biomedical Research (NIBR), and performed under the co-supervision of the professor Erik Van Nimwegen (*Computational & Systems Biology*, Biozentrum) at Basel university.

The aim of this project was to develop and apply innovative bioinformatics methods to toxicogenomic data generated mainly from IMI-MARCAR consortium in order to gain a better understanding of the early gene regulatory processes underlying non-genotoxic carcinogenesis in the context of drug safety assessment.

This thesis is organized as follows. In **Chapter 2** we first introduce the problem of non-genotoxic carcinogenesis in the context of drug safety assessment. We then briefly present the liver and discuss important mechanisms of hepatocarcinogenesis along with experimental models with a focus on Phenobarbital-promoted liver tumor rodent model. We finally give an overview of toxicogenomic data and bioinformatic approaches to model transcriptional regulatory networks. The main findings of the thesis, that are arranged in two manuscripts, are then each covered in the central chapters of this thesis. **Chapter 3** shows how adapting existing probabilistic algorithm to comprehensive toxicogenomic data from *in vivo* experiments leads to identification of key regulatory interactions underlying early stages of drug-induced liver tumorigenesis. This manuscript has been published in *Nucleic Acid Research* journal in January 2014. **Chapter 4** describes a study where human relevance of rodent humanized model is discussed in terms of gene expression data. This manuscript has been published in *Toxicological Sciences* in April 2014. Of note only the material that was considered sensible for complete publication in peer reviewed journals is reported in this thesis. The thesis concludes by a discussion on the major findings, their implications for drug safety assessment, an outlook of where future work could be taken up and the remaining open questions (**Chapter 5**).

Contents

Chapter 1

Introduction

Non-genotoxic carcinogens (NGC) form a group of molecules that do not directly bind DNA (1) but that produce perturbations in the gene expression and epigenetic state of cells (2; 3; 4) which facilitate tumor formation, typically through the promotion of preexisting neoplastic cells into neoplasms (5; 6). The molecular events underlying the NGC-induced transformation of normal hepatocytes to altered hepatocellular foci are still unclear and no acute early molecular markers for NGC are available for drugs under development. The significant delay in drug development due to positive findings of drug-induced non-genotoxic carcinogenesis together with the fact that many environmental pollutants, industrial chemicals, and food contaminants are potential NGC that have not been adequately tested for carcinogenicity are some of the reasons that motivate toxicologists to develop early biomarkers of NGC and improve early safety assessment of such compounds.

According to regulatory expectations, drug safety is tested in both short term *in vitro* and long-term *in vivo* studies in several experimental animals (rodent and non-rodent species) prior testing on human (7). As the safety assessment in experimental animals has been very successful in predicting toxicity of biologically active chemicals in humans (8), differences in species biochemistry or pathophysiology between human and rodents have raised doubts regarding the appropriateness of extrapolating some rodent tumor findings to humans (9). A better understanding of NGC mode of action on cellular mechanism is believed to help addressing the relevance of rodent assays to human risk assessment (10; 9) and help in early prediction of NGC in drug development.

Toxicogenomics is a ten years old discipline that applies genomic science to toxicology. It allows to investigate the molecular and cellular effects of chemicals in biological systems and thus complements biochemical and phenotypic classic approaches leading to both drug toxicity and drug mode of action identification. Furthermore toxicogenomic data are particularly suitable for early biomarkers as genomic perturbations are often detectable prior to phenotypic symptoms.

In this dissertation we have adapted innovative bioinformatic approaches to toxicogenomic data from comprehensive *in vivo* experiments in order 1) to identify key early regulatory interactions underlying liver drug-induced non-genotoxic carcinogenesis and 2) to examine potential species-specificity (human-mice) in receptor-dependent mechanisms underlying liver tissue molecular responses to NGC. The outcome of this research provides with novel mechanism-based candidate biomarkers for NGC, and allows for a better understanding of early mechanisms and pathways underlying drug-induced toxicity in rodents and their relevance to human.

Chapter 2

Background

2.1 Safety assessment in drug in development

2.1.1 Preclinical safety assessment

The approximately 15-years long process of drug development comprises the assessment of the drug efficacy, bioavailability and safety (11). Preclinical safety assessment is a crucial step that takes place early on during the drug development process. It is intended to i) define the target organ toxicity of the tested compound, ii) estimate a safety margin between the efficacious dose and the dose causing an adverse effect, iii) predict drug toxicity in humans and iv) eventually identify a maximum recommended safe starting dose (MRSD) (12) for the clinic as reviewed in (13). Importantly ≈30% of failures in the development of drugs are related to toxicity and safety issues as reviewed elsewhere (11) making it a serious impediment to development of new medicines. Drug safety is assessed in both short term *in vitro* and long-term *in vivo* studies in several experimental animals (rodent and non-rodent species) before testing on human as required by regulatory agencies (7).

2.1.2 Hepatotoxicity

One of the major safety issue in drug development is hepatotoxicity due to the facts that 1) the liver has the greatest biotransformation capability for the processing of chemicals and thus is involved in the metabolism of nearly all xenobiotics; and 2) that the liver is exposed to the largest amounts of chemicals absorbed from the gastrointestinal tract (14). Consequently the liver is a primary target organ for most chemicals irrespective of their mode of action (14). Importantly liver adapts to drug exposure in a way that is not necessarily toxic. Indeed the three following types of morphologic alterations of the liver can occur upon xenobiotic exposure depending on dose and duration (as reviewed in (15)): (i) adaptive alteration that consists of an exaggerated normal physiologic response; (ii) pharmacologic alteration, that consists of an expected alteration in response to the desired action of the test compound; and (iii) adverse alteration that consists of morphologic alterations that are generally undesired, progressive and deleterious to the normal function of the cell(s) involved.

Hepatotoxicity can primarily result from 1) inhibition of mitochondrial function, 2) disruption of intracellular calcium homeostasis, 3) activation of apoptosis, 4) oxidative stress (16), 5) inhibition of specific enzymes and transporters, and 6) formation of reactive metabolites that cause direct toxicity or immunogenicity (17). Hepatic adaptive liver response upon chemical exposure often results in enhanced tissue capacity to dispose of the chemical, via for example induction of phase I and II enzymes that catalyze bio-

transformation of the inducing chemical. Adaptive response usually leads to changes in gene expression, alteration of the metabolome and increase in liver size (18) that are reversible upon cessation of exposure, preserve viability and are not considered toxic (14). Hepatocytes hypertrophy can also be observed that reflects hyperplasia of organelles, mainly endoplasmic reticulum and peroxisomes as increase in functional demand leads to organelle expansion and thus enhancing the capacity of the liver to respond to stress.

An adaptive effect can however become adverse if exposure exceeds a certain threshold (that can be time or dose related) leading to disruption of equilibrium, compromised tissue viability and, in the worse case, liver tumor (14). These are typically induced upon the prolonged creation of reactive oxygen species (19; 20; 21), or covalent binding with cellular macromolecules (22). Of note biliary system, hepatic vasculature, Kupffer cells, or stellate cells (Ito cells) can also be targeted by tested compounds and involved in adverse effects (15).

2.1.3 Carcinogenicity testing and non-genotoxic carcinogens

The most expensive, time- and animal-consuming test in preclinical safety aims to identify chemicals that may pose potential human carcinogenic risk compared to the benefit for the therapeutic indication (23). Carcinogenicity testing is required prior to registration of many new pharmaceutical agents intended for chronic or intermittent use over 6 months of duration (24), and clinical considerations include the expected duration of treatment, the severity of the disease or disorder, the nature and size of the patient population, and the availability of other therapies as reviewed in (25).

Carcinogenic compounds are classified either as genotoxic or non-genotoxic carcinogens (NGC) depending on whether their carcinogenicity resides or not in their ability to interact with DNA and induce DNA mutation and repair responses. Genotoxic carcinogens induce structural DNA changes leading to pro-carcinogenic mutations and can be organized according to their structural features such as alkenes, aromatic amines and nitrosamines. Conversely NGC, initially designated as “epigenetic” carcinogens by Weisburger and Williams (1981) (26), are non DNA-reactive compounds that produce epigenetic effects on cells, that either indirectly result in DNA modification or facilitate development of preexisting neoplastic cells into neoplasms (27).

Genotoxic carcinogens are inexpensively identified in the early stage of drug development using *in vitro* assays (28; 29). There is however currently no sufficiently accurate and well-validated short-term assay to identify NGC and NGC identification largely relies on 2-year rodent bioassays which current protocol involves exposing a large number of animals (50-70 male and female rats and mice per group) to varying doses of the studied chemical with histopathological assessment of multiple organs and tissues in each of the animals at the end of the 2-year exposure period as reviewed in (24; 25). Importantly, as this test is time-consuming, labor-intensive long and costs millions of dollars per compound, it is often planned late in the development process (30). The identification of early mechanisms-based biomarkers for NGC would therefore allow for the design of more predictive tests that would eventually lead to significant improvement in cancer risk assessment of compound in development.

As mentioned previously, the liver is the major target organ of chemically induced toxicity and the most prevalent drug-induced tumor site in both male and female mice and rats according to the National Toxicology Program (NTP) database and the Carcinogenic Potency Database (CPD) (25), and as such a leading single cause for withdrawal of approved drugs from the U.S. market (15; 17). This thesis focuses on NGC-induced liver tumorigenesis and the following sections briefly introduces liver physiology.

2.2 Liver physiology

The liver is a vital organ that has a pivotal role in human body metabolic homeostasis. Liver functions include but are not limited to i) glucostat activity i.e glucose release and production via glycogenolysis and gluconeogenesis respectively (31), ii) bile acid formation, iii) filtering activity of the blood coming from the digestive tract, iv) metabolic homeostasis of carbohydrates, amino acids, lipids and lipoproteins and v) detoxification of numerous endo- and exogenous substances (32). Accordingly the liver is highly responsive to environmental perturbations such as changes in portal blood composition (33).

Oxygenated blood from aorta enters the liver through the hepatic artery. Nutrients enriched blood containing immune complexes and xenobiotics arrives from gastrointestinal tract, spleen and pancreas and enters the liver via the portal vein; it then proceeds through the sinusoids (surrounded by a single cell layer consisting of about 20 hepatocytes) and eventually drains into the central venule located at the center of each lobule, the microscopic functional unit of the liver (34). Liver also produces bile that is transported away to larger bile ducts via bile ductule (inverse flow direction as blood, see **Figure 2.1**). Portal vein, hepatic artery, and bile ductule compose the portal triad (34).

Most liver functions are endorsed by the hepatocytes, that constitute the major cellular compartment of the liver. Hepatocytes are aligned on plates of one cell thick as depicted in **Figure 2.1** extending from the portal triad in linear fashion to the central vein, with two basolateral domains facing the sinusoidal space from which uptake of blood-borne contents takes place as reviewed in (14). Together with portal triad and central veinule they form the microscopic functional unit of the liver tissue designated as the hepatic lobule (32) (see **Figure 2.1**). Hepatocytes are connected via gap junctions formed by connexons allowing fast cell-cell communication between adjacent hepatocytes (14).

Portal blood is progressively filtered by hepatocytes and a decreasing gradient of nutrient and oxygen is created from periportal to perivenous regions. Pathologists commonly discern 3 zones (see **Figure 2.1**) in liver lobule that follows the bloodstream: the periportal region perfused with blood rich in oxygen, substrates and hormones (zone 1), the perivenous region, that receives blood with low oxygen content (zone 3), and the zone in between (zone 2). Liver zonation is also reflected by differences in hepatocyte ultrastructure that correlate with different enzymatic activities and gene expression. Periportal hepatocytes have larger and fewer mitochondria (32) and are specialized in oxidative energy metabolism, amino acid catabolism, ureagenesis, gluconeogenesis, cholesterol synthesis and selected types of protective metabolism as reviewed in (35). Conversely perivenous hepatocytes have more abundant endoplasmic reticulum, express most CYP forms and perform preferentially glycolysis, glycogen synthesis from glucose, liponeogenesis, glutamine formation, and xenobiotic metabolism (35).

Hepatocytes occupy almost 80% of the total liver volume and also perform the majority of liver functions; 10% of the liver volume (and 40% to the total number of liver cells) is occupied by sinusoidal endothelial cells (SEC), Kupffer cells (resident liver macrophages), hepatic stellate cells (fat- and retinoids-storing cells) and pit cells (large granular lymphocytes) generally being more numerous in the periportal region (32; 36). As the majority of liver functions are carried by hepatocytes, these cells are also the main targets of liver damaging agents.

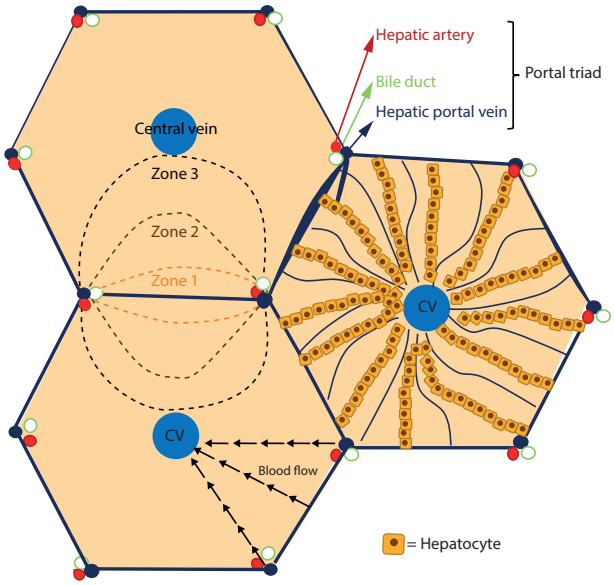

Figure 2.1: **Basic architecture of the liver lobule.** Oxygenated blood from aorta and nutrient enriched blood from gastrointestinal tract enter the liver through the hepatic artery and portal vein respectively. Portal blood then proceeds through the sinusoids, surrounded by a single hepatocyte layer, until the central venule. Bile is transported away to larger bile ducts via bile ductule. Portal vein, hepatic artery, and bile ductule compose the portal triad. Differences in hepatocyte ultrastructure that correlate with different enzymatic activities and gene expression discern 3 zones in liver lobule that follows the bloodstream: the zone 1 or periportal region, the zone 2 or midzonal region and the zone 3 or pericentral region.

2.2.1 Liver proliferation

In normal adult liver, less than 5% hepatocytes undergo proliferation; this reflects a low rate of cell death through apoptosis (37). The liver has however a substantial regenerative capacity that is reflected by the complete recovery of the liver upon partial resection or severe injury. This phenomenon results from rapid proliferation of all the existing mature cellular populations composing the intact organ to restore organ mass (32) and does not necessarily depend on progenitor or stem cells (34). In this process the regenerative response is tightly regulated to be proportional to the amount of liver removed and to result in a liver size proportional to the body size (38).

The regulation of hepatocyte proliferation has been subjected to extensive investigations (see (39) for review) and while the exact mechanisms responsible for the exit from the quiescent state and the re-entry into the cell cycle remain unclear, sequential changes in gene expression, growth factor production, and morphologic structure have been shown to take place during this process (34). Extracellular factors and paracrine interactions with neighboring non-parenchymal liver cells such as Kupffer and Ito cells have been moreover shown to be essential components of this machinery (39). Interestingly mitogenic response upon liver injury has been shown to occur in different population of hepatocytes (originating from different zones) according to the type of stimuli i.e. type of chemical exposure, and reduction in liver mass (39; 40; 41; 42).

2.2.2 Liver polyploidy

Progressive nuclear polyploidization occurs widely in metabolically active tissues and is a characteristic feature of mammalian hepatocytes that takes place during postnatal growth (43). About 70% of adult hepatocytes in rodents and 40% in humans are tetraploid (44; 45; 46). Polyploidy in hepatocytes is initiated in postnatal liver growth and can result from different mechanisms (**Figure 2.2**) that include i) incomplete cytokinesis (leading to binuclear polyploid hepatocytes), ii) endoreplication, defined as cycles of DNA replication in the absence of mitosis or iii) endomitosis where mitosis is interrupted (47; 48; 49; 46; 50). Thus polyploid hepatocytes can be either mononuclear or binuclear. Sister chromatids in polyploid cells are associated either with a single centromere or have distinct centromeres for all of their chromosomes depending on whether their result from endocycling or from endomitosis (43; 51).

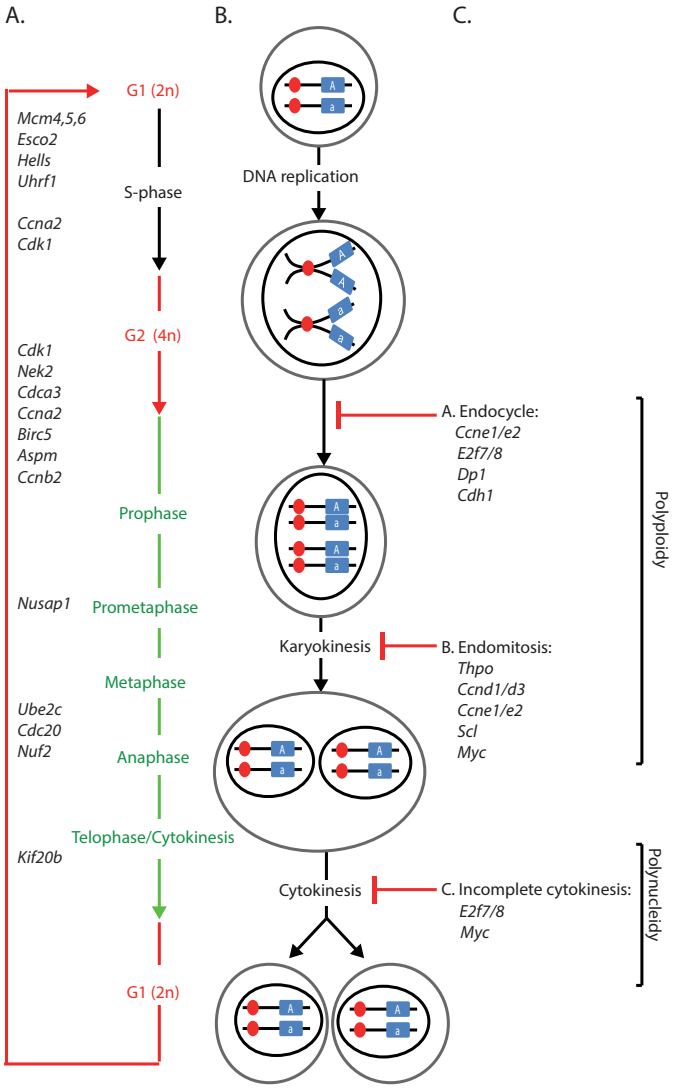

Figure 2.2: **Mechanisms leading to hepatocyte polyploidy| A.** Schematic representation of different cell cycle stages with genes that are differentially expressed upon PB treatment between day 1 and day 3 and involved in regulating these stages (see **Chapter 4**). **B.** Schematic representation of DNA content along the different cell cycle stages. **C.** Polyploidy or polynucleidy can result from incomplete cytokinesis, endocycle or endomitosis that are likely regulated by genes reviewed in Pandit et al, (2013) (52).

The regulatory mechanisms underlying polyploidization are not completely clear, however regulation of mitosis and cytokinesis have been identified as key processes. Insulin signaling and downstream regulation of the PI3K/Akt signaling pathway that controls cytoskeleton organization has been related to cytokinesis failure (50; 53; 54). More recently E2F7 and E2F8 were shown to inhibit the completion of cell division thus enhancing hepatocytes polyploidy and binucleation in liver development and regeneration, whereas the canonical activator E2F1 was shown to counteract their activities (55; 56). A consequence of the increase in cellular DNA content is an increase in cellular volume that was demonstrated in studies with both human and mouse liver cells where the volume of hepatocytes was approximately twice with doubling DNA content (57; 58; 59; 60).

As liver polyploidy is not necessary for the liver to fulfill its functions, the role of increase in ploidy in the liver is not entirely clear. Some speculate that endoreplication occurs as a mean to increase the availability of DNA copies and thus increase gene expression (43). As oxidative liver damage has been associated with a pronounced increase in the population of polyploid cells, and ligands of nuclear receptors such as PB and TCPOBOP have been shown to cause liver polyploidisation (39; 61; 62), polyploidisation was proposed as a mean to increase resistance to genotoxic damage and apoptosis (63).

While hepatocyte polyploidy generally occurs in cells that are terminally differentiated (43), liver tumor lesions such as hepatocarcinoma, hepatoadenoma and early liver lesions (see **Section 2.3** for terminology of liver tumors) are characterised by lower polyploid fraction compared to an age-matched normal liver in both humans and carcinogen-induced rodent models (64; 44; 65; 66; 67). While some propose that selective proliferation of mononucleated 2n hepatocytes could be one of the early events of the liver transformation process and thus proposing polyploidization as a tumor-suppressor function, others argue that polyploidization being linked to chromosomal instability might promote tumor development (68).

2.3 Liver tumorigenesis and Hepatocarcinoma (HCC)

As mentioned earlier, observed neoplastic lesions following long-term exposure to both genotoxic and non-genotoxic chemicals are predominantly liver tumors arising from hepatocytes (69) and are therefore a key area in drug safety. Hepatocytes-derived liver tumors can start as hepatocellular adenoma (HCA) that are benign liver tumors composed of non invasive multilayered differentiated hepatic plates (70). HCA are usually well demarcated as they show prominent compression of the surrounding tissues (71). HCA can in rare cases transform into hepatocellular carcinoma (HCC) (also named hepatoma), the most frequent malignant liver cancer (72). HCC can be well differentiated lesions or undifferentiated cells, have undefined borders and are diffusively infiltrative cancer.

Development of hepatocellular carcinoma (HCC) is a complex and long process that involves hepatocyte transformation in neoplastic cells, inhibition of apoptosis, stimulated angiogenesis, reprogramming of energy metabolism, evasion to immune destruction and invasion in surrounding tissues via tissue remodeling; these are key features of HCC that are mostly shared among any cancers, as extensively reviewed in (73). In the case of human HCC, these features are often the consequence of chronic inflammation (as a result of liver cirrhosis and chronic hepatitis) and subsequent liver fibrosis (74; 75) that is the fifth cause of cancer death worldwide. Dramatic changes in gene expression accompany all mechanisms associated with HCC development, from the transformation of normal hepatocyte into neoplastic cells to the establishment of a tumor-prone environment (76). Importantly as liver tumors have been well described and characterized, the underlying gene regulatory mechanisms bridging the long-term effect of chronic inflammation or drug exposure to the hallmarks of cancer in HCC remain largely unknown.

In the following we review some key aspects that significantly influence liver tumor development with a focus on gene expression and transcription factors, key regulators of gene expression.

2.3.1 Gene regulatory mechanisms in HCC development

2.3.1.1 Gene expression regulation

Gene expression regulation is a complex cellular process that is summarized in **Figure 2.3**. Data and papers generated by the Encyclopedia of DNA Elements (ENCODE) consortium in September 2012 largely contribute to the current knowledge of functional elements in the human genome sequence (77; 78; 79; 80) and point towards a higher complexity of gene regulation than was previously believed.

Gene expression is first regulated at transcriptional level (step 1 in **Figure 2.3**). Transcriptional regulation is a key and complex mechanism that depends on the presence of a specific combination of transcription factors (TFs) and co-factors in both the promoter regions of genes and in regulatory sites located more distant from the genes (leading to DNA looping and long-range interactions (81)), that altogether facilitate RNA Polymerase II recruitment and binding to upstream gene promoters and eventually determine the onset and rate of RNA synthesis (see (82) for review on transcriptional regulation); RNA Polymerase II is indeed responsible for transcribing protein-coding genes and miRNA. While numerous proteins such as chormatin remodellers, polymerase and helicase are involved in regulating transcription, DNA binding TFs play central role in this mechanism as they bind to specific DNA sequences of promoter and distal regions (80; 83; 84) also designated as transcription factor binding sites (TFBS) (79; 85; 86; 87); the specific combination of TFs contained in the regulatory regions then eventually determines which specific subset of genes is expressed under which condition. This mechanism is particularly crucial for the cell to fulfill its function in appropriate time and condition as it allows complex and precise patterns of the expression of the 40,000 genes contained in human genome with the $\approx$ 1,900 human TFs (88) thus enabling the cell to respond to intrinsic and extrinsic cues such as drug-induced response in case of hepatocytes. Importantly the TFs DNA binding rate (step 8 in **Figure 2.3**) also depends on 1) nuclear concentration in TFs and co-factors (step 7 in **Figure 2.3**), and 2) the local cell-dependent chromatin context (79; 80; 89; 90) such as histone modifications, nucleosome positioning and DNA methylation; DNA methylation indeed defines feature of mammalian cellular identity (91) and is itself influenced by DNA-binding factors, especially in Low Methylated Regions (LMRs) where the presence of DNA-binding factors and their binding is necessary and sufficient to determine the low methylation status of these regions (92). As a consequence, DNA methylation pattern highly correlates with global occupancy patterns of major sequence-specific regulatory factors (93).

RNA post-processing, that includes RNA splicing (94) and polyadenylation, and subcellular localization (78) are additional regulatory mechanisms of gene expression, that depend on a complex machinery of RNA binding proteins and interactions with several RNA molecules (miRNAs and lncRNA) that eventually determine mRNA stability and degradation (steps 2 and 3 in **Figure 2.3**). Importantly DNA methylation and GC architecture have also been shown to regulate RNA splicing (95; 96).

mRNA is eventually destined to migrate to cytoplasm where protein translation can start (step 3 and 4 in **Figure 2.3**). Post-translational modifications then determines protein activation and localization (steps 5 and 6 in **Figure 2.3**). If the protein encodes a TF, nuclear translocation and interaction with co-factors (steps 7 and 8 in **Figure 2.3**) eventually lead to gene expression. Importantly all of the regulatory steps described in **Figure 2.3** can be disrupted in cancer leading to aberrant protein expression and disrupted cell functions.

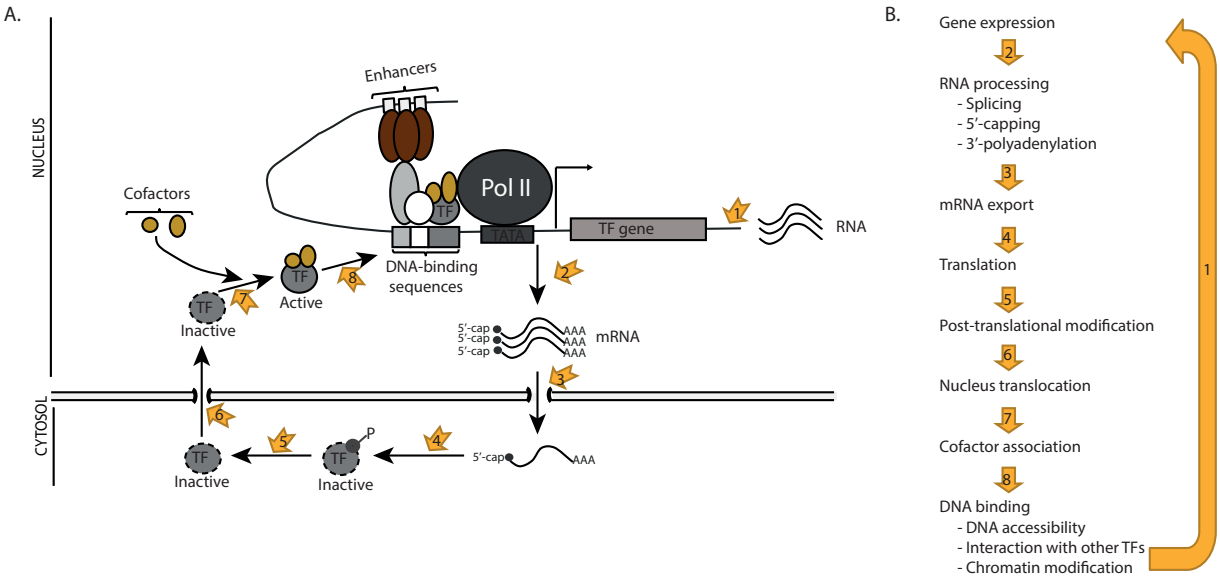

Figure 2.3: **Gene expression regulation| A)** Gene expression regulation encompasses several levels from gene transcription to DNA binding. **B)** Overview of 8 steps of gene expression regulation.

2.3.1.2 Genetic mutations in HCC development: genotoxic carcinogens MOA

Accumulation of genetic mutations is a widely accepted cellular process that initiates the slow transformation of a normal cell into neoplastic cells, especially mutations in oncogenes and tumor-suppressor genes (73; 97) and the carcinogenicity of genotoxic compounds resides in their ability to interact with DNA and induce DNA mutation. Genetic mutations can affect 1) regulatory regions, leading to aberrant transcriptional rate, 2) intron-exon boundary regions, leading to aberrant splicing, or 3) exonic regions, leading to change in protein conformation that can affect protein stability and protein interaction with partners and DNA that can then impact cancer development. Recent progresses and easier access to DNA-sequencing technologies allowed for establishment of exhaustive list of genetic mutations in HCC (98; 99; 100).

2.3.1.3 Epigenetic changes in HCC development: non-genotoxic carcinogens MOA

With the advent of genome-scale methylome and histone marks characterization, epigenetic disruptions emerged as an additional fundamental basis for cancer initiation and progression that contribute to stabilization of pre-existing genetic mutations and/or activation/silencing of oncogenes and tumor-suppressor genes via changes in chromatin status of their regulatory regions (97; 101). Epigenetic modifications are not restricted to chromatin status modifications and encompass disruption in any of the regulatory steps mentioned in **Figure 2.3** including TFs and cofactors changes in activity, and disruption in splicing machinery. NGC carcinogenicity resides in their ability to stabilize pre-neoplastic mutated cells and promote their growth via such epigenetic changes. As reviewed earlier, while genotoxic compound are easily identified early on during process of drug development, NGC class of compound poses major issues for preclinical toxicity testing as their mode of action is more complex (longer time-scale of disease and requirement of interaction between different cellular compartments). There is therefore a great need for increasing understanding of underlying regulatory mechanisms.

2.3.1.4 Transcription factors in liver non-genotoxic carcinogenesis

TFs are considered as key intrinsic regulators of mechanisms underlying epigenetic reprogramming associated with cancer development (102). As reviewed in (103) activation or aberrant expression of TFs frequently represents the last step in a number of signaling pathways that affect proliferation, apoptosis, migration, or senescence in an oncogenic manner (103). Indeed aberrant TFs activity has been frequently associated with human (104; 105; 106) and mouse (107) HCC, and numerous TFs have been shown to play central role in regulatory mechanisms of liver cancers. FOXA1/FOXA2 for example were shown to regulate molecular mechanisms responsible for gender dimorphism in HCC (107). Activation of the β-catenin pathway in hepatoadenoma (108; 109; 110) and in 20% to 34% of hepatocellular carcinomas (111; 112; 113; 114) suggest important regulatory role for β-catenin in HCC.

Several intracellular pathways can modulate TFs activity including disrupted or facilitated DNA binding efficiency through modification of chromatin status and DNA accessibility, changes in TF nuclear concentration, and changes in nuclear concentration of co-factors. Changes in RNA processing and splice variants can also modulate DNA binding affinity associated with change in TFs activity (115; 116).

NGC compounds such as phenobarbital (PB) induce progressive chromatin remodeling and changes in gene expression in target tissue of carcinogenicity (117; 118; 119; 120; 3; 121; 122; 4; 123). Furthermore some of these changes have been shown to target key drivers of cell proliferation such as *Fos* and *Myc* (124). All these perturbations involve TFs change in activity, and thus the identification of TFs participating to regulation of all stages of non-genotoxic carcinogenesis is a crucial step towards assessing carcinogenic potential of novel therapeutics and improving the understanding of their MOA. Methods to identify dysregulated TFs are discussed in **Section 2.5**.

2.3.2 Hormonal perturbation in HCC development

Lower spontaneous liver tumor incidence is observed in females as compared to males in both humans (125) and rodents (126) strongly supporting a key role of hormones in liver cancer development (127). This gender disparity in liver cancer was shown to result in part from estrogen-mediated inhibition of IL-6 expression in Kupffer cells that, in turn, was shown to affect hepatocyte proliferation (128). More strikingly female mice deficient for *Esr1* lost their resistance to HCC (128). Additional studies support a liver tumor promoter role of androgens (via induction of DNA damage and oxidative stress) (129) and a liver tumor suppressor role of estrogens (through reduction in the proinflammatory effects of MyD88-mediated secretion of IL6) (128).

Thus hormonal perturbation is considered as a key mode of action in non-genotoxic carcinogenesis (9). Several NGCs are indeed hormonally active agents such as the anti-epileptic phenobarbital (PB), the best characterized NGC in rodents (please see **Section 2.4.3** for detailed description of PB-mediated liver tumor promotion), that has been shown to induce very quick response of the pituitary gland in both humans (130; 131) and rodents (132). However while a sexually dimorphic regulation of phenobarbital-induced cytochromes P450 2B1 and 2B2 has been shown in rat (133) that may be responsible for the gender-specific regulation of xenobiotic-induced hepatocyte proliferation in mice (134), sexual dimorphism in chemically-induced liver tumorigenesis is more controversial. Results from few studies performed in both males and females rats or mice over long period (more than 2 years) suggest that, while PB promotional effect is similar in males and females i.e. equal increase in tumor occurrence, female mice live longer under PB-treatment and promotional effect in female require more time (135; 136; 137). In this thesis we focussed on male mice, however we are aware that our results and the conclusions that we draw from

these studies can significantly differ in female context.

2.3.3 Microenvironment in HCC development

Microenvironment plays a key role in liver tumor development (see (138) for review) and more generally in cancer development. The liver microenvironment mainly consists in endothelial cells, Kupffer cells and stellate cells, that are involved in immune response and fibrosis, but also in cytokines, growth factors and several proteins. Constant communication between hepatocytes, stellate cells and Kupffer cells via cytokines secretion contribute indeed to create a tumor-prone environment. For example, preneoplastic hepatocytes, that have high proliferating potential, do not necessarily grow autonomously, and communication with non-parenchymal cells via paracrine factors has been shown to play key role in this process (71). Furthermore hepatic stellate cells (Ito cells), that are activated upon liver injury, seem to play dual role: as on one hand they produce excess extracellular matrix (ECM) and thus participate in ECM remodelling (71), on the second hand they seem to communicate with macrophages to modulate liver fibrosis upon liver chronic injury (139). Extracellular matrix remodeling is an additional key process in HCC development that sustains hepatocytes proliferation by providing cells with a reservoir for a variety of cytokines and growth factors (76).

2.4 Rodent models of HCC

In order to study the mechanistic and cellular aspects of liver tumor biology including genetics of tumor initiation and promotion, *in vivo* tumor progression and spreading (metastasis), rodent experimental models of HCC remain the standard (140). Indeed rodent models are the only available assay that allows non-genotoxic hepatic carcinogenicity assessment of compounds in development (27; 141). Of note *in vitro* cultured hepatocytes are valuable tools for hepatotoxicity testing, but their use is limited to short-term studies due to rapid reduction in cytochrome P450 (CYP) activities caused by a decrease in CYP transcription and an alteration in the expression of key transcription factors when cultured on plastic (142).

During the last decades numerous experimental models of chronic or acute liver induced carcinogenesis have been developed as reviewed in (140). The two-stage experimental models are often used that comprise an *initiation phase*, during which short-term exposure to a genotoxic compound induces genetic alterations, followed by the *tumor promotion phase*, during which long-term exposure to non-genotoxic compound accelerates the outgrowth of pre-existing mutated cells and thus the process of tumor development.

2.4.1 Tumor initiation and genotoxic carcinogens

Rodent models of liver carcinogenesis are often initiated with carcinogenic compound that induce random genetic mutations and therefore accelerates tumor occurrence. The most widely used experimental model is the diethylnitrosamine (DEN)-induced liver carcinogenesis (143). A single dose of DEN at the age of 2 weeks causes DNA-damage in mice leading to HCC at approximately 8-10 months of age (140). As reviewed in (140) the carcinogenic capacity of DEN resides in its capability of alkylating DNA structures but also in the oxidative stress caused by DEN (144; 145). The carcinogenic potential of DEN as well as the time needed after a DEN-injection to develop HCC does depend on the administered dose, the sex, the age and the strain of mice (140; 146).

2.4.2 Tumor promotion and non-genotoxic carcinogens

Tumor promotion is defined as a process through which pre-neoplastic or initiated cells evolve into a malignant neoplasm under the action of exogenous or endogenous compound that fix pre-existing mutations. The study of the biological processes underlying tumor promotion is of great importance for both 1) human cancer research, as several endogenous molecules have been shown to promote tumors (147; 148), and 2) drug development process, as tumor promotion is the MOA of NGC that are difficult to identify in early phases of drug development.

Importantly characterization of all stages of tumor promotion process is necessary as, contrarily to irreversible effects of genotoxic agents, tumor promotion by NGC is reversible until a certain stage (149). However because tumor promotion develops over a long time period and involves several cellular compartments, the regulatory mechanisms underlying all stages of tumor promotion are poorly characterized.

Several rodent models of non-genotoxic HCC have been developed, both endogenous model of liver carcinogenesis such as the methyl-deficient model (150; 151), as well as exogenous models of liver carcinogenesis, for example methapyrilene (histamine receptor antagonis), diethylstilbestrol (DES, an estrogen receptor agonist), Wy-14643 (Wy, a peroxisome proliferator activated receptor α agonist), piperonyl-butoxide (PBO, a pesticide synergist) (152). In the next section we focus on the best characterised exogenous tumor promoter, the phenobarbital, that is the model used in this thesis.

2.4.3 Phenobarbital (PB)

The most widely used anticonvulsant phenobarbital (PB) is a well established rodent NGC used to investigate the promotion of non-genotoxic HCC in rodent livers (153; 154; 136). PB functions as a tumor promoter by increasing the incidence of spontaneously and chemically induced tumors (155; 153; 154; 136; 156). As reviewed earlier, PB induces progressive chromatin remodeling and changes in gene expression in target tissue of carcinogenicity, the liver (117; 118; 119; 120; 3; 121; 122; 4; 123). Importantly, although liver tumors only develop after 35 weeks of chronic exposure to PB, changes in gene expression and chromatin modifications are detected as soon as one day after treatment initiation with PB, particularly in genes involed in drug metabolism and xenobitic response such as *Cyp2b10* (4; 121; 122; 123).

PB accomplishes its diverse effects on liver function in part by promoting nuclear translocation of the constitutive androstane receptor (CAR) (157) which reflects both acute and chronic response to PB treatment (158; 159; 160). When PB is used as a tumor promoter subsequent to DEN, more than 80% of liver tumors harbors activating mutations in β-catenin (161) which prevents the phosphorylation of β-catenin by the Axin/ APC/CK1/GSK3 complex (162; 163; 164), and thereby the subsequent degradation of β-catenin by the proteosome (165). This leads to enhanced translocation to the nucleus (166) resulting in the aberrant interaction with a variety of transcription factors and subsequent activation of target genes (167; 168; 169). Conversely in absence of tumor-promoting agents, mouse liver tumors are frequently mutated in Ha-ras, mutation otherwise undetectable in promoted tumors, whilst Ctnnb1 mutations are almost absent liver tumors induced by DEN alone (161; 170; 171; 172).

Importantly PB-mediated promotion effect on DEN-initiated mice varies depending on strain, sex and age of the mice as reviewed in (140). PB generally promotes liver tumor after DEN initiation, however a tumor inhibiting effect has been observed in B6C3F1 male mice when exposure to DEN was performed in the pre-weaning stage (173; 174); this effect was however absent from female B6C3F1 mice (175) and from male Balb/c (176) and C3H (177) mice.

Several mechanisms of PB-mediated liver tumor promotion have been proposed that include induction of oxidative stress upon PB-mediated increased cytochrome P450 activity (178) and hypermethylation in promoter regions of tumor suppressor genes (118). However these mechanisms remain hypotheses and whilst a recent study identified *Meg3* as an early candidate biomarker for NGC in rodents (123), underlying regulatory mechanisms of PB-mediated tumor promotion remain elusive. In the following two sections we present two previously identified regulators involved in PB-mediated tumor promotion.

2.4.3.1 Constitutive Androstane Receptor

PB-mediated CAR nuclear translocation is a critical process which induces both acute and chronic response to PB treatment, and is required for gene expression changes, hepatomegaly and liver tumor formation elicited upon prolonged PB treatment in mice (159; 160). CAR is a member of the nuclear steroid and thyroid hormone receptor superfamily but, unlike classic nuclear hormone receptors which are activated by their cognate ligands, CAR is a transcription factor that is indirectly activated by various xenobiotics, and is transcriptionally active in the absence of exogenous hormone (179). CAR is involved in several key processes of liver physiology such as drug metabolism, hepatic energy metabolism, cell growth, and cell death (180; 181; 182; 183).

In non-induced mice, CAR is phosphorylated at Thr38 by signaling induced by epidermal growth factor (EGF) (184), and forms a complex with heat shock protein 90 (HSP90) that prevents its nuclear translocation (157; 185). Hepatocytes exposure to PB inhibits EGFR signaling (186), leading to dephosphorylation of cytoplasmic CAR upon protein phosphatase 2A recruitment to the CAR:HSP90 complex, that facilitates CAR nuclear translocation. Thus PB-induced CAR nuclear translocation is regulated through cascade of phosphorylation-dephosphorylation (157; 185; 187).

CAR transient activation induces hepatic expression of detoxification enzymes and transporters, and transient hepatomegaly (158; 188; 156) that augments the ability of the liver to metabolizes PB. Conversely chronic PB-mediated CAR activation induces complex dynamics of transcriptional response (123) and chromatin remodeling (3; 4; 121; 122), hepatocytes hypertrophy and accelerates development of liver tumors (158). Whilst CAR is essential for liver tumorigenesis in response to chronic treatment with PB (189) it is not necessary for liver hepatocarcinogenesis as demonstrated by similar tumor prevalence in non-treated CAR null and wild-type mice (189).

Please note that in addition to CAR, PB also activates the pregnane X receptor (PXR) (190), which has overlapping functions with CAR to coregulate xenobiotic metabolism and detoxification in liver (181; 191), and whose co-activation may enhance CAR-mediated hepatocyte proliferation (192).

2.4.3.2 β-catenin

Wnt/β-catenin pathway plays key roles in liver physiology including liver organogenesis (please see (193) for review), metabolic zonation of adult liver (194), hepatocytes proliferation and liver regeneration, epithelial-mesenchymal transition, and cell adhesion (by its association with epithelial cadherin and actin) (195; 196). β-catenin is also required for metabolism of ammonia (197) and is involved in the regulation of inducible expression of P450s and drug-metabolizing enzymes mediated by xenobiotic-sensing receptors such as Ahr and CAR (198; 199; 200; 201).

β-catenin regulates gene expression through the growth factor-β-catenin/T-cell factor-4 (TCF4) signal-

ing pathway. Furthermore β-catenin core domain contains 12 armadillo repeats which drive interaction with additional nuclear TFs and target genes (202). The nuclear amount of β-catenin is regulated by phosphorylation of its N-terminal region as reviewed in (71): cytoplasmic phosphorylation of β-catenin by glycogen synthase kinase-3b (GSK3b) leads to rapid protein ubiquitination and subsequent degradation, that can only occur when β-catenin forms a complex with the adenomatosis polyposis coli (APC), Axin and GSK3b. Upon Wnt binding to Frizzled, Dishevelled is phosphorylated leading to inhibition of GSK3β and thus allowing for accumulation of β-catenin, nuclear translocation and target gene activation such as *cyclin D1*, *c-myc*, *PPAR*δ, *Hnf1*α and *CD44* as reviewed in (169).

As mentioned earlier, long-term PB treatment stimulates clonal expansion of a dormant initiated cell population mutated in β-catenin and represses clonal expansion of H-ras mutated cells that display hepatocarcinogenicity in absence of PB; it is however noteworthy that in the absence of PB treatment, *Ctnnb1* mutations are almost absent of mouse liver tumors induced by DEN (161; 170; 171; 172). Hepatocytes bearing mutation in β-catenin display increase in de-phospho β-catenin and enhanced translocation to the nucleus (161; 167; 168; 169; 203; 204). Approximately 30% of human HCC and 15% of hepatic adenoma have β-catenin activating mutations suggesting that the protein plays key role in HCC development (111; 112; 113; 114; 205) in both humans and rodents. The role of β-catenin activation in HCC development is however not clear and several studies demonstrate that β-catenin activation alone is not sufficient for HCC development. Indeed truncated mutation in N-terminal region of the protein is not enough to provide with proliferative advantage in absence of PB treatment (206) and Wnt pathway activation by stabilized β-catenin was shown to be insufficient for hepatocarcinogenesis (207); the latter observation is in line with observed low prevalence of *Ctnnb1* mutated tumors after DEN-initiation in absence of PB treatment. Moreover β-catenin has been shown to prevent tumor development in absence of PB treatment by restricting oxidative stress, inflammation and fibrosis (208; 209). Thus additional (epi)-genetic alterations must be involved in β-catenin activated HCC development.

Long-term PB treatment is apparently leading to proper (epi)-genetic alterations as it results in outgrowth of β-catenin activated hepatocytes specifically, whilst preventing outgrowth of H-ras mutated hepatocytes. Importantly β-catenin knock-out (KO) animals are completely resistant to PB-mediated liver tumorigenesis. Underlying regulatory mechanisms responsible for the outgrowth of *Ctnnb1* mutated hepatocytes upon long-term PB exposure remain however largely unknown. The hypothesis that PB may select for β-catenin activated hepatocytes by interfering with β-catenin/(LEF/TCF)- dependent transcriptional programs was rejected by Aydinilik *et al*, (2001) (161) due to the fact that both β-catenin and Cyclin D1 protein levels remained equally elevated in promoted and non-promoted tumors. Furthermore direct PB-mediated Wnt/β-catenin pathway activation is not supported by the absence of liver-specific β-catenin target gene (such as glutamine synthetase) up-regulation upon PB treatment.

In conclusion while β-catenin is necessary for PB-mediated tumor promotion and its activation is a hallmark of PB promoted liver tumors, its role in the pathogenesis remains elusive. Current studies converge on the idea that β-catenin mutation is necessary but insufficient on its own for HCC and requires cooperation with additional regulators and pathways to results in unrestricted hepatocyte proliferation as reviewed in (39).

2.4.3.3 Remaining open question

As reviewed in precedent sections, β-catenin and CAR are two known regulators involved in PB-mediated tumor promotion. However while crucial roles for both proteins in PB-mediated rodent liver tumor

promotion have been demonstrated, none of them seems to be sufficient for the process. As β-catenin is not directly activated by PB, progressive selection for β-catenin activated hepatocytes upon PB chronic exposure is more likely to result from interaction with alternative activated cellular pathway. Finally while CAR has been shown to be a key regulator of PB-induced gene expression, its constant activation upon PB treatment cannot explain the complex dynamics of transcriptional response observed in the first 3 months of PB treatment (123). In conclusion additional transcription factors must be involved in this process and their identification holds great promise in increasing understanding of PB-promoted liver tumorigenesis. Several methods exist to identify additional transcription factors involved in this process, especially from gene expression data; these are reviewed in **Section 2.5** and **Chapter 3** presents our innovative computational approach that led to the identification of new candidate regulators of PB-mediated tumor promotion.

2.4.4 Human relevance of rodent model of HCC

The safety assessment in experimental animals of biologically active chemicals has been very successful in predicting toxicity in humans (8). However differences in species biochemistry, pathophysiology, or drug pharamacology between human and rodents have raised doubts regarding the appropriateness of extrapolating some rodent tumor findings to humans (9; 210). Indeed whilst prolonged treatment with PB does increase liver size in humans (211), human hepatocytes are resistant to the ability of PB to increase hepatocyte proliferation (212; 213) and inhibit apoptosis (213). Consequently PB-induced rodent non-genotoxic hepatocarcinogenesis is not considered to be a relevant mechanism for humans (9) and there is no evidence of a specific role of PB in human liver cancer risk based on epidemiological data in epileptics (214; 215; 216).

Humanized mouse models for drug metabolizing enzymes and to a lesser extent drug transporters in which the endogenous mouse genes have been replaced with human genes have been used in drug development to explore the species specificity of drug toxicity and to overcome the limitation of animal models in accurately predicting human responses (217; 218; 219; 220; 221). These include humanised mouse models in which the endogenous mouse CAR/PXR genes have been replaced with human CAR/PXR genes (222). However mouse genetic context largely differs from that of human (including difference in co-factors, chromatin status, promoter of target genes, TFBS) and apart from direct drug-mediated activation and target genes investigations in murine context, their relevance to human response is still controversial. Alternatives to these are chimeric mice, which have human hepatocytes engrafted in their liver and that have been used to study human drug metabolism and pharmacodynamic responses for nearly 20 years (223).

2.5 Regulatory mechanisms investigations in biological systems

The rapid evolution of genomic-based technologies led to the emergence of toxicogenomics (224) defined as the application of genomic science to toxicology. This approach allows to improve the understanding of the molecular and cellular effects of chemicals in biological systems and thus complements biochemical and phenotypic approaches in assessing the toxicology of a compound.

Gene expression data generated from DNA microarrays or RNA sequencing have been the most successful type of data used in toxicogenomics. As many biological responses to xenobiotics are manifest at the transcriptional level (nuclear receptor activation induced upon drug exposure is implicitly followed by changes in gene expression (225)), gene expression data from *in vivo* and *in vitro* models have been

used to 1) delineate mechanisms of compound toxicity and 2) identify predictive molecular markers of toxicity by studying the function of the affected genes (24; 30; 226; 227; 228; 229; 230; 231). Furthermore building up databases of gene expression changes associated with various toxic compound exposure can be used to discriminate toxic from benign response (232) and to classify drugs (30).

The identification of drug-induced differentially expressed genes holds great promise for establishment of early biomarkers of drug toxicity. Furthermore the identification of the regulators responsible for the observed changes in gene expression -that eventually lead to perturbation of biological pathways and cellular states- provide with a mechanistic understanding of therapeutics MOA and thus is a crucial step towards assessing carcinogenic potential of novel therapeutics. Numerous methods - computational and experimental - have been developed towards identifying and validating candidate regulators of biological processes that are reviewed below.

2.5.1 Computationally-based methods to identify dysregulated TFs

Because RNA represents the direct output of TFs activity in the cells, a wealth of computational systems biology studies are focussed on developing methods to reconstruct transcriptional regulatory networks from gene expression data and identify TFs that regulate and determine the context-specific expression of a gene (233).

Classic approach to predict TF activity Differentially expressed TFs identified either with gene expression data or with Reverse Protein Arrays (RPA) are often the first approach to predict key dysregulated TFs. However because TFs activity is regulated at several level (expression, translation, PTMs, cellular localization, interaction with co-factors and DNA binding) (234), TF activity does not necessarily correlate with concentration or expression level (235) and alternative methods are required to predict dysregulated TFs.

Cluster of co-expressed genes Classic methods for modeling transcriptional regulatory networks from gene expression aim at collecting genes in co-expressed clusters (236). Numerous relevance scores have been proposed to cluster genes in modules such as correlation coefficient score (237), mutual information (238; 239; 240), and singular value decomposition. Then, assuming that co-expressed genes are co-regulated by a common set of TFs, the corresponding regulatory regions of each genes in the cluster can be extracted and over-representated TFBS are then considered candidate common regulatory elements for these clusters (87; 86; 241). Numerous sequence analysis approaches have been developed which identify potential TF binding sites in DNA sequences set (85). These methods are however prone to significant noise as many of the predicted potential TF binding sites are not functional (242). Furthermore these methods are limited by the need to detect motif influence from statistically aggregated expression data rather than from individual genes and this typically restricts their application to subsets of genes with large gene expression signals as reviewed in (243).

Combining gene expression experiments with TF-gene network topology A major challenge in reconstructing transcriptional regulatory networks resides in the fact that as one TF may control the expression of up to hundreds of genes, one gene is often regulated by a combination of TFs and miRNAs as reviewed in (244). Consequently methods that model explicitly genome-wide gene expression patterns in terms of condition-specific TFs post-translational activities and gene-specific regulatory network connectivity (see (236) for review) embrace this aspect. Gene-TF connectivity can be obtained from databases such as RegulonDB (234), using chromatin immunoprecipitation data (245; 246; 247), from

analysis of promoter region, or by identifying which genes are differentially expressed when the TF is deleted (248; 249). It is however important to note that the structure of the regulatory network of the cell can change dramatically between different experimental or environmental conditions (250).

Several methods have been proposed to solve linear model that links condition-specific TFs activities to gene expression of target genes and gene-specific regulatory network connectivity. Da *et al,* (2006) propose to use line spline functions to correlate the binding strengths of motifs with the expression levels (251). Gao *et al,* (2004) developed an algorithm that predicts TFs activities based on ChIP and transcriptome data using multivariate regression and backward variable selection (246). Nguyen *et al,* (2006) uses a deterministic mathematical strategy for deriving principles of transcription regulation at the single-gene resolution level (243). Finally Suzuki *et al,* (2009) uses a bayesian framework to solve the multivariate linear model (252). This method is further developed in the following section. Importantly predicting TFs activities from gene expression is likely to provide with more accurate prediction than ChIP experiments as some studies revealed that there is little overlap between the genes whose promoters are bound by a TF and those whose expression changes when the TF is deleted (253). In general, the genes whose promoters are bound by a TF according to ChIP-chip experiments and those whose expression level responds to perturbation of the same TF show little overlap - typically 3-5% (254).

Motif Activity Response Analysis (MARA) As mentioned in previous paragraph, Suzuki *et al,* (2009) uses a bayesian framework, Motif Activity Response Analysis (MARA), to solve the multivariate linear model (252). MARA models gene expression dynamics explicitly in terms of predicted number of functional TFBS N_{pm} within proximal promoter regions (-300 to +100) of the genes and the post-translational activities of their cognate transcription factors. The model assumes that the expression e_{ps} of a promoter p in sample s is a linear function of the activities A_{ms} of motifs m that have predicted sites in p such as:

$$e_{ps} = \tilde{c}_s + c_p + \sum_m N_{pm} A_{ms}$$

where c_p reflects the basal activity of promoters p and $\tilde{c}_s$ reflects the total expression in sample s. The number of functional TFBS N_{pm} are predicted using the Bayesian regulatory-site prediction algorithm MotEvo that incorporates information from orthologous sequences in six other mammals and uses explicit models for the evolution of regulatory sites (255). As a result, MARA provides for a total of 189 TFBS motifs (that represent the DNA binding specificities of close to 350 TFs) the activity profiles of these regulators across the samples, the significance of each motif in explaining the observed expression variation across the samples, their target genes, and the sites on the genome through which these regulators act. The activity A_{ms} of a motif m in a sample s represents the condition dependent nuclear activity of positive and negative regulatory factors that bind to the sites of the motif. As motif activity is inferred from the behavior of the predicted targets of the motif, an increasing activity is inferred when its predicted targets show on average an increase in expression, that cannot be explained by the presence of other motifs in their promoters. The details of the method are described elsewhere (252; 256).

Importantly instead of predicting gene expression from TFs activities and regulatory regions contained in their promoters, this algorithm aims at predicting key regulators that drive gene expression changes across the samples, their activities across the samples, and their genome-wide targets. Inferring regulatory activities from the behavior of predicted targets instead of the expression profile of the TFs themselves allows for prediction of differential activity that are not related to the expression of the TF but rather due to post-translational modifications, changes in cellular localization, or interactions with co-factors.

The output of MARA is a concrete set of hypotheses that are readily amenable to direct experimental validation and follow-up.

However this method is currently limited by the TFBS data-base as in mammals sequence-specificities are available for only about 350 of the about 1,500 TFs. Another limitation of the MARA algorithm is that it focuses solely on predicted TFBSs in proximal promoters, ignoring the effects of distal enhancers. Moreover motifs mode of action cannot be distinguished and therefore TFs which are both activator and repressor under a certain condition will not be detected.

2.5.2 Experimentally-based approaches to validate predicted dysregulated TFs

Various experimental methods intended to validate predicted dysregulated TFs have been successfully applied that comprise knock-in (KI) (160; 220), knock-out (KO) or siRNA gene silencing experiments (3; 129; 159; 197; 208; 257; 252; 249). In these experiments TFs coding genes are either inserted, removed or silenced from experimental models allowing to demonstrate the causal relationship between the regulator and a given phenotype; these methods do not however allow the identification of the direct target genes. Conversely chromatin immunoprecipitation (ChIP) experiments which identify all regions bound by a given DNA-binding regulators provide lists of direct TFs target genes together with their genomic context (79; 80; 258; 233). Furthermore comparisons of ChIP binding results from different experimental conditions allow to investigate differential TF activity between conditions. ChIP experiments are however limited in the number of tested hypotheses due to poor antibody quality and ChIP-based activity prediction can yield low accuracy as TF binding does not necessarily imply a regulatory function. Immunohistochemistry staining experiments complement the above approaches and are often utilized to 1) precisely identify which cell population in a tissue expresses a given TF and 2) investigate tissue heterogeneity in TF expression. Additional *in situ* hybridization of target mRNA and co-localization with TF allows furthermore to demonstrate direct physical interaction between a TF and a target gene of interest.

2.6 IMI-MARCAR

Most gene expression data used in this thesis have been generated throughout IMI-MARCAR consortium (http://www.imi-marcar.eu/). The MARCAR project is a 5 year project funded under the Innovative Medicines Initiative (IMI) Joint Undertaking that aims at (i) identify early biomarkers for more reliably predicting which compounds have a potential for later cancer development, (ii) improve the scientific basis for assessing carcinogenic potential of non-genotoxic drugs, (iii) identify the molecular response to NGC exposure that underpins development of early exposure biomarkers and finally (iv) improve drug safety and the efficiency of drug development by progressing the development of alternative research methods.

Chapter 3

Computational modeling identifies key gene regulatory interactions underlying phenobarbital-mediated tumor promotion

This chapter contains the main manuscript of this thesis which shows how adapting existing probabilistic algorithm to comprehensive toxicogenomic data from *in vivo* experiments leads to identification of key regulatory interactions underlying early stage of drug-induced liver tumorigenesis. The manuscript has been published in *Nucleic Acid Research* in January 2014.

Nucleic Acids Research Advance Access published January 23, 2014

Nucleic Acids Research, 2014, 1–16
doi:10.1093/nar/gkt1415

Computational modeling identifies key gene regulatory interactions underlying phenobarbital-mediated tumor promotion

Raphaëlle Luisier[1], Elif B. Unterberger[2], Jay I. Goodman[3], Michael Schwarz[2], Jonathan Moggs[1], Rémi Terranova[1,*] and Erik van Nimwegen[4,*]

[1]Discovery and Investigative Safety, Novartis Institutes for Biomedical Research, 4057 Basel, Switzerland, [2]Department of Toxicology, Institute of Experimental and Clinical Pharmacology and Toxicology, University of Tübingen, 72074 Tübingen, Germany, [3]Department of Pharmacology and Toxicology, Michigan State University, MI 48824, USA and [4]Biozentrum, University of Basel and Swiss Institute of Bioinformatics, 4056 Basel, Switzerland

Received September 25, 2013; Revised December 20, 2013; Accepted December 24, 2013

ABSTRACT

Gene regulatory interactions underlying the early stages of non-genotoxic carcinogenesis are poorly understood. Here, we have identified key candidate regulators of phenobarbital (PB)-mediated mouse liver tumorigenesis, a well-characterized model of non-genotoxic carcinogenesis, by applying a new computational modeling approach to a comprehensive collection of *in vivo* gene expression studies. We have combined our previously developed motif activity response analysis (MARA), which models gene expression patterns in terms of computationally predicted transcription factor binding sites with singular value decomposition (SVD) of the inferred motif activities, to disentangle the roles that different transcriptional regulators play in specific biological pathways of tumor promotion. Furthermore, transgenic mouse models enabled us to identify which of these regulatory activities was downstream of constitutive androstane receptor and β-catenin signaling, both crucial components of PB-mediated liver tumorigenesis. We propose novel roles for E2F and ZFP161 in PB-mediated hepatocyte proliferation and suggest that PB-mediated suppression of ESR1 activity contributes to the development of a tumor-prone environment. Our study shows that combining MARA with SVD allows for automated identification of independent transcription regulatory programs within a complex *in vivo* tissue environment and provides novel mechanistic insights into PB-mediated hepatocarcinogenesis.

INTRODUCTION

Aberrant activity of transcription factors (TFs) is a hallmark of both human (1–3) and mouse (4) hepatocarcinogenesis and is considered as a key intrinsic regulatory mechanism underlying epigenetic reprograming associated with cancer development (5). Non-genotoxic carcinogens (NGC) are a group of compounds that do not directly affect DNA (6), but that produce perturbations in the gene expression and epigenetic state of cells (7–9) which, if given in sufficient concentration and duration, facilitate tumor formation, typically through the promotion of pre-existing neoplastic cells into neoplasms (10,11). However, little is known about the regulatory mechanisms that underly the tumor promotion by NGC, particularly regarding the early regulatory changes in response to the carcinogen.

The anticonvulsant phenobarbital (PB) is a well-established rodent NGC that has been extensively used to investigate the promotion of liver tumors (12–14). PB accomplishes its diverse effects on liver function, at least in part, by promoting nuclear translocation of the constitutive androstane receptor (CAR) (15) through inhibition of Epidermal Growth Factor Receptor (EGFR) signaling (16). CAR activation is required for the acute and the chronic response to PB treatment and for liver tumor formation elicited upon prolonged PB treatment (17–21). In addition to this crucial role of CAR, when liver tumors are promoted through PB treatment in combination with an initial treatment with diethylnitrosamine (DEN), >80% of the resulting tumors harbor activating mutations in β-catenin (22) that stabilize β-catenin, leading to enhanced nuclear translocation and subsequent target gene activation (23–27).

Apart from the crucial roles for CAR and β-catenin in PB-mediated liver tumor promotion, little is known about

*To whom correspondence should be addressed. Email: erik.vannimwegen@unibas.ch
*Correspondence may also be addressed to Rémi Terranova. Email: remi.terranova@novartis.com

additional transcriptional regulators that orchestrate the complex and dynamic PB-mediated gene expression programs associated with early molecular responses to PB treatment (28) and long-term PB tumorigenic effects (12–14).

In this study, we have elucidated gene regulatory interactions underlying dynamic PB-mediated transcriptional responses during the early stages of liver non-genotoxic carcinogenesis by integrating multiple gene expression datasets from independent *in vivo* mouse PB studies. Our primary dataset consists of an early kinetic study (seven time points across 91 days of PB treatment) originally designed to investigate the temporal sequence of molecular and histopathological perturbations during the early stages of PB-mediated liver tumor promotion *in vivo* (9, 28). Several challenges are associated with the extraction of key gene regulatory interactions from gene expression time course data. First, we needed to identify the relative contributions (activities) of specific transcriptional regulators underlying the observed genome-wide gene expression changes. This was achieved using our recently developed motif activity response analysis [MARA, (29)]. MARA capitalizes on sophisticated computational methods, developed over the last decade (30), that allow comprehensive prediction of binding sites for hundreds of mammalian TFs across all mammalian promoters (31). Using such computational predictions, MARA models observed gene expression patterns explicitly in terms of the predicted regulatory sites and uses this to infer the regulatory activities of TFs. A number of recent studies (32–43) demonstrate that this approach can successfully identify key regulators *ab initio* across different model systems of interest.

A second challenge was to disentangle the complex range of PB-mediated gene expression programs in mouse liver tissue that are associated with distinct biological events including xenobiotic responses, tumor promotion and tumorigenesis.

Here, we show that combining MARA with singular value decomposition (SVD) allows for automated disentangling of independent transcription regulatory programs within a complex *in vivo* tissue environment. We were able to successfully infer key gene regulatory proteins for xenobiotic responses, tumor promotion and end-stage tumors as well as assess their genetic dependence on CAR and β-catenin signaling pathways.

Collectively, our analyses provide novel mechanistic insights into PB-mediated tumor promotion in the mouse liver, including a proposed role of E2F and ZFP161 in regulating PB-mediated hepatocyte proliferation at both early and tumor stages and progressive PB-mediated suppression of ESR1 activity that likely contributes to the development of a tumor-prone environment.

MATERIALS AND METHODS

Gene expression datasets and Affymetrix GeneChip processing

A library of 109 genome-wide messenger RNA (mRNA) expression patterns was compiled from four different studies (Figure 1a). In all four studies gene expression was profiled using Affymetrix GeneChip MOE-4302 (Affymetrix, Santa Clara, CA, USA). The analysis of the micro-array data was done with the R statistical package, version 2.13 (2005) and Bioconductor libraries, version 1.4.7 (44).

From gene expression matrices to motif activity matrices

Matrices of activities for 189 mammalian regulatory motifs across all samples were inferred from the RMA-normalized expression matrices using the MARA algorithm (29) (Figure 1b and c). MARA models genome-wide gene expression patterns in terms of predicted functional Transcription Factor Binding Sites (TFBSs) within proximal promoter regions (running from −300 to +100 relative to transcription start) of the 40 300 promoters. The model assumes that the expression e_{ps} of a promoter p in sample s is a linear function of the predicted numbers of binding sites N_{pm} for each motif m in promoter p and the (unknown) activities A_{ms} of each of the motifs m in sample s, i.e.

$$e_{ps} = \tilde{c}_s + c_p + \sum_m N_{pm} A_{ms}$$

where c_p reflects the basal activity of promoter p and $\tilde{c}_s$ is a normalization constant corresponding to the total expression in sample s. The activities A_{ms}, as well as error bars δA_{ms} on these activities, are thus inferred from the measured expression data e_{ps} and the predicted binding sites N_{pm}. The number of functional TFBSs N_{pm} was predicted using the Bayesian regulatory site prediction algorithm MotEvo, which incorporates information from orthologous sequences in six other mammals and uses explicit models for the evolution of regulatory sites (30). The 189 regulatory motifs represent binding specificities of roughly 350 different mouse TFs. Besides the motif activities, MARA also calculates a z-score quantifying the significance of each motif in explaining the observed expression variation across the samples, the target genes of each motif, and the sites on the genome through which the regulators act on their targets.

Formally, the activity A_{ms} corresponds to the amount by which the expression e_{ps} would be reduced if a binding site for motif m in promoter p were to be removed. Thus, an increasing activity is inferred when its targets show on average an increase in expression, that cannot be explained by the presence of other motifs in their promoters. The details of the method are described elsewhere (29). An overview of the analysis strategy and an outline of the MARA approach are depicted in Figure 1.

Detection of differential motif activity between pairs of conditions

We quantified the differential motif activity between two conditions using a z-statistic as

$$z_{m\Delta c} = \frac{\bar{A}_{mc_1} - \bar{A}_{mc_2}}{\sqrt{\delta A^2_{mc_1} + \delta A^2_{mc_2}}}$$

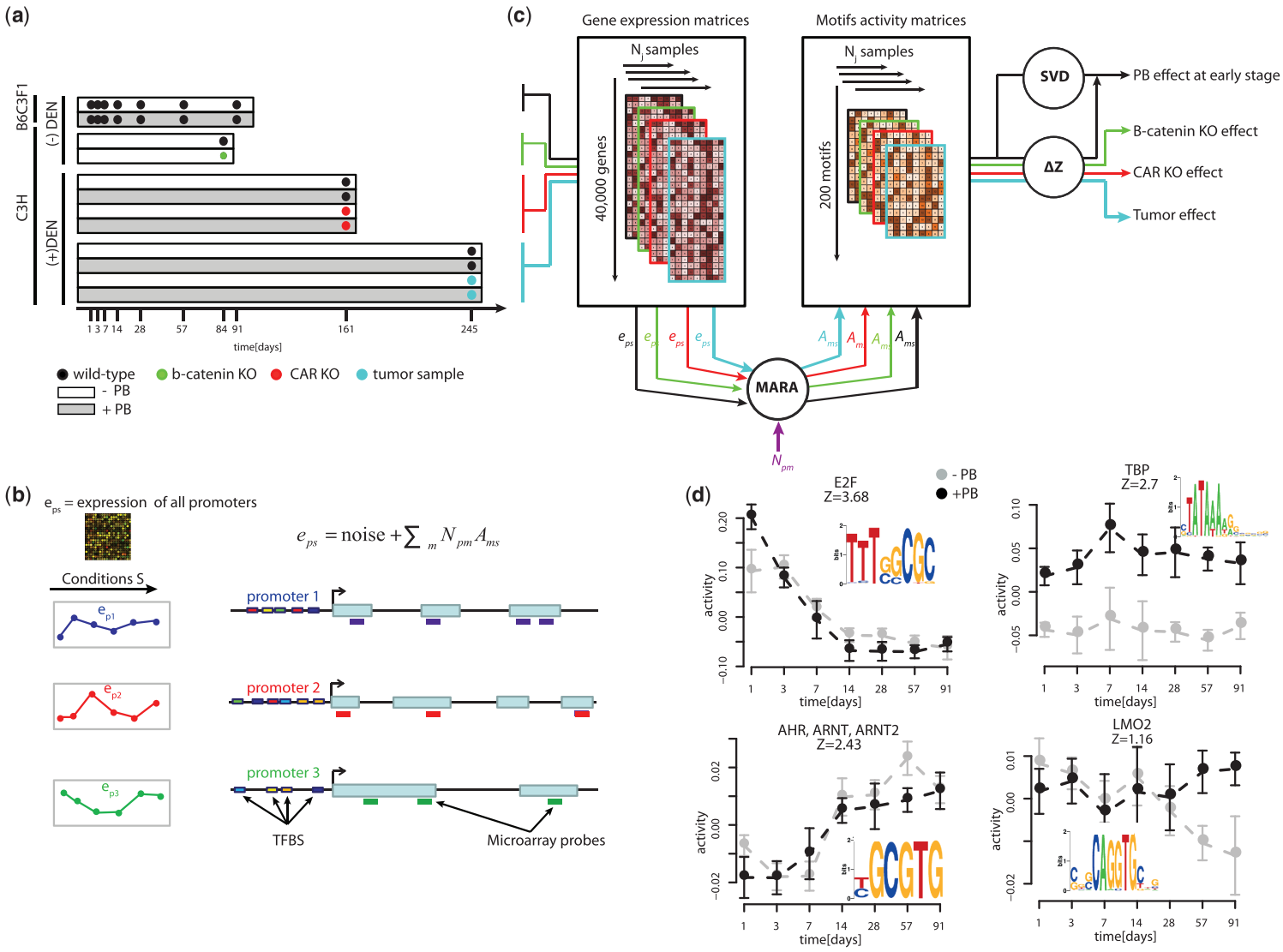

Figure 1. Overview of the analysis strategy for predicting TF activities from gene expression data. (**a**) Schematic representation of the four toxicogenomic datasets used in this study. Each row corresponds to an independent experimental treatment with gray indicating PB treatment and white indicating control treatment. Time is indicated along the horizontal axis on the bottom and the black dots indicate the times at which samples were taken for transcriptomic profiling. Colored dots indicate whether the study involved WT animals (black), liver-conditional β-catenin null mice (green), CAR null mice (red) or tumor cells (blue). The mouse strain used in each of the studies as well as the presence or absence of DEN treatment is indicated on the left-hand side. Multiple biological replicates were performed for each of the studies. (**b**) Outline of the MARA approach. Using measured intensities of micro-array probes together with known gene structures, MARA first estimates the log-expression e_{ps} from each promoter p in each sample s. Second, MARA makes uses of comprehensive computational predictions of transcription binding sites in mammalian promoters, with N_{pm} denoting the total number of predicted regulatory sites for motif m in promoter p. MARA then models the expression levels e_{ps} as a linear function of the numbers of computationally predicted binding sites N_{pm} and unknown 'motif activities' A_{ms} of each motif m in each sample s. That is, the algorithm uses the linear model to infer the motif activities A_{ms}. (**c**) Overview of the computational analysis applied to our combination of datasets. First, MARA is applied to all expression datasets and motif activities A_{ms} are inferred for all motifs across all samples. These motif activities are then further subjected to SVD analysis and analysis of various contrasts (Δz). (**d**) Examples of inferred motif activities along the time 91-day time course. Each panel corresponds to one motif, with the name of the TF indicated on top of the panel and its sequence logo as an inset. Black and gray lines are predicted activities in PB-treated and control samples, respectively.

where $\bar{A}_{mc}$ is the averaged motif activity profile over replicates for condition c and motif m and δA_{mc} is the standard error on the corresponding motif activity, which is computed using a rigorous Bayesian procedure (Balwierz, PJ. et al, manuscript under review). The z-values quantify the evidence for a change in regulatory activity of the motif between the two conditions. That is, if $z_{m\Delta c}$ is highly positive it indicates that predicted targets of motif m are upregulated in condition c_1 relative to condition c_2, in a way that cannot be explained by the activities of other regulators. We consider motifs differentially active if $|z| \geq 1.5$.

In order to avoid any confounding batch effects, we only calculate differential activities across conditions from the same dataset. Comparing activities between treated and control samples at different timepoints of the kinetic study allowed for the identification of PB-mediated dysregulated TFs at the early stage of PB treatment. Comparison of activities between wild-type (WT) and CAR/β-catenin null in physiological conditions, i.e. without treatment, identified motifs whose activities are modulated upon KO of the respective TF (Figure 2a and c). We consider such motifs to be downstream of the β-catenin/CAR pathways in physiological conditions. Similarly, comparison of activities between PB-treated and non-treated samples allowed for the identification of motifs that are dysregulated by PB treatment. By further comparing the changes in motif activities upon PB treatment for both WT and CAR-null samples, we can identify motifs dysregulated by PB in a manner that is independent of CAR signaling and motifs whose dysregulation is downstream of CAR (Figure 2a and c). Comparison of activities between promoted tumors and surrounding PB-treated tissue identified motifs dysregulated in promoted tumors; comparison of activities between non-promoted tumors and surrounding non-treated tissue identified motifs dysregulated in liver tumors irrespective of PB treatment. Motifs uniquely dysregulated in promoted tumors were classified as promoted tumor-specific regulators (Figure 2b).

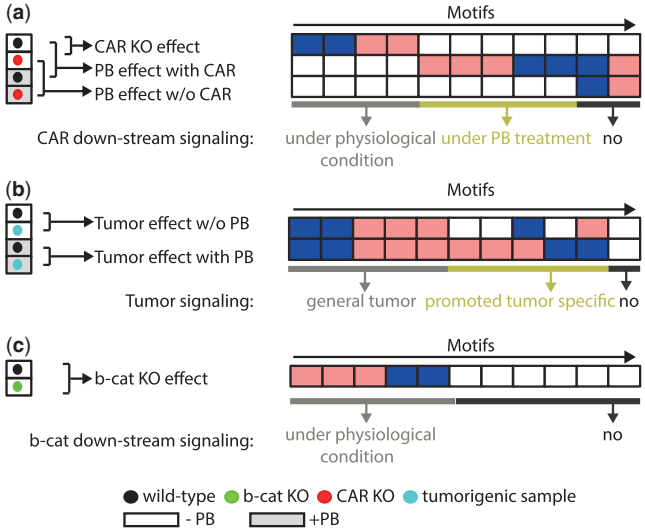

Figure 2. Schematic representation of contrasts applied for differential motif activity analysis of each dataset. The color of the dot indicates whether the sample is WT (black), β-catenin KO (green), CAR KO (red) or tumor (blue). White boxes correspond to control samples and gray boxes to PB-treated samples. The arrows show which pairs of samples are compared for each contrast and point to corresponding rows with example motif activity changes (blue corresponding to downregulation $z < -1.5$, pink to upregulation $z > 1.5$ and white no significant change $|z| < 1.5$). (**a**) Motif activities from the CAR KO study are compared to identify regulators downstream of the CAR pathway under physiological conditions and under PB treatment. (**b**) Motif activities from the tumor study are compared to identify promoted tumor-specific regulators. (**c**) Motif activities from the β-catenin KO study are compared to identify downstream regulators of the β-catenin pathway under physiological conditions.

Characterization of PB-mediated early motif activity profiles

SVD of the motif activities

We performed SVD of the activities of the 189 motifs across the seven timepoints in PB- and vehicle-treated livers, i.e. a matrix A containing 189 rows and 14 columns. SVD resulted in a decomposition, $A = U\Lambda V$, where Λ is a diagonal matrix containing the singular values, U and V contain the orthonormal bases defined by right and left singular vectors of A, respectively. Each motif activity profile $\vec{a}_m$ with $(\vec{a}_m)_s = A_{ms}$ can be thought of as a linear combination of the right singular vectors $\{\vec{v}_k\}$.

Visualization and interpretation of the SVD results

To visualize the right singular vectors $\{\vec{v}_k\}$, we plotted the activities v_{ks} on the vertical axis as a function of the time corresponding to each sample s on the horizontal axis and coloring all samples corresponding to PB treatment black, and those corresponding to control-treatment gray, e.g. Figure 1d. This visualization facilitated the biological interpretation of the singular vectors. Biological interpretation was further facilitated by identification of the regulatory motifs whose activity profiles correlate most strongly (either positively or negatively) with the activity profile of the singular vector.

Identification of representative motifs of the singular vectors

As the right singular vectors form an orthonormal basis of the space of activity profiles, the projection of a given motif activity profile onto a right singular vector indicates how strongly the motif's activity profile overlaps with the basis vector specified by the singular vector. The projections of the motif activity profiles $\vec{a}_m$ onto right singular vectors $\vec{v}_k$ are calculated as $q_{mk} = \vec{a}_m \cdot \vec{v}_k$ and these values are readily obtained from the SVD results as $AV = U\Lambda$ such that $q_{mk} = (U\Lambda)_{mk}$.

We additionally computed Pearson correlations between the motif activity profiles $\vec{a}_m$ and the right singular vectors $\vec{v}_k$. As the vectors $\vec{v}_k$ are linear combinations of the motif activity profiles $\vec{a}_m$ that are mean centered, i.e. $\sum_s a_{ms} = 0$, these are also mean centered. Consequently, the Pearson correlation coefficients can also be readily obtained from the SVD results as

$\rho_{mk} = q_{mk}/\sqrt{\sum_{k'}(q_{mk'})^2}$. As the activity profiles of different motifs have different overall 'lengths', the projections and Pearson correlations do not carry identical information. Motifs with large activities tend to have high absolute projections with a given singular vector, even if the motif activity profile is not similar to the activity profile of the singular vector. In contrast, a motif with small activities will tend to have low projections, but may have a high correlation with a given singular vector.

In order to identify representative motifs for each singular vector, motifs were ranked according to both projection and correlation scores. The highest (most positive scores in both projection and correlation) and lowest (most negative scores in both correlation and projection) motifs were selected for each singular vector. As some degree of redundancy is present among regulatory motifs, we further refined our motifs selection in a systematic manner following criteria that are detailed in the 'Results' section.

Gene Ontology enrichment analysis

The DAVID Bioinformatics Resource (Database for Annotation, Visualization and Integrated Discovery) (45,46), version 6.7, sponsored by the National Institute of Allergy and Infectious Diseases (NIAID), NIH, was used to investigate the statistical enrichment of biological terms and processes associated with the predicted target genes of each motif of interest. We directly imported official gene symbols into DAVID, exported enrichment from biological pathways from Gene Ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG), filtered out redundant terms and selected biological processes with P-value of enrichment <0.05.

RESULTS

Overview of liver toxicogenomic data from phenobarbital-treated mouse models

In order to investigate gene regulatory networks underlying early PB-mediated liver tumor promotion, we used four transcriptomic datasets which are illustrated in Figure 1a. Our primary dataset is composed of transcriptome profiling data from a PB kinetic study in B6C3F1 (livers from vehicle, i.e. control and PB-treated male mice at +1, +3, +7, +14, +28, +57 and +91 days of dosing). This dataset enabled us to investigate gene expression dynamics during the first 3 months of PB treatment. A second CAR knock-out (KO) study composed of transcriptome profiling data from livers of vehicle- and PB-treated C3H male WT and CAR-null mice (at +161 days of dosing) enabled us to investigate which of the responses to PB treatment were CAR-dependent at this later time point. A third tumor study consisting of samples from promoted (at +35 weeks of PB treatment) and non-promoted tumors as well as their related surrounding tissue from C3H male mice, enabled us to identify gene regulatory changes that were specific to promoted tumors, as opposed to being a shared feature of tumor tissues in general. Finally, a β-catenin KO study composed of livers from WT and β-catenin-null C3H male mice enabled us to investigate which of the identified TFs were downstream of β-catenin in physiological conditions. In both the CAR KO and tumor studies, mice were DEN-initiated at 4 weeks of age.

Identifying PB-modulated activities of transcriptional regulators using MARA

MARA is a general method for inferring the activities of a large collection of mammalian TFs (as represented by their DNA binding 'motifs') by modeling gene expression data in terms of computationally predicted regulatory sites in promoters. The basic approach is illustrated in Figure 1b. Note that motif activities are inferred from the behavior of the expression levels, typically hundreds, of predicted 'targets' of the motif and do not directly involve analysis of the expression levels of the regulators themselves. This is especially useful in systems where TF activities are modulated through subcellular localization and post-translational modifications, rather than at the transcriptional level, e.g. such as the PB-mediated CAR nuclear translocation and induction of downstream transcriptional responses that we study here. Importantly, apart from inferring the motif activities A_{ms}, MARA also rigorously infers error bars on these motif activities δA_{ms}, which allow to quantify to what extent motif activities are significantly varying across the samples for each motif. The overall significance of each motif m is then represented by a z-statistic ('Materials and Methods' section).

TFs underlying early PB-mediated liver transcriptional dynamics

Figure 1d shows the activities of four motifs observed within the time course of control and PB-treated mice, illustrating the range of different profiles that can be observed. For example, the motif bound by the family of E2F TFs and the motif bound by AHR, ARNT and ARNT2 TFs both showed substantial changes in activity across the time course that are largely the same in the control and PB-treated animals, except for E2F's activity at the first timepoint. In contrast, the TATA-box motif bound by TATA Binding Protein (TBP) exhibited almost constant activity across time but showed a strong shift in behavior between control and PB-treated animals. The LMO2 motif showed no significant activity for the first month of the time course but at later time points (during the last 2 months) there was a marked divergence between PB-treated and control animals.

SVD identifies four characteristic motif activity profiles underlying early PB-mediated transcriptional changes

Although it is possible to formulate biological interpretations and hypotheses for observed motif activity profiles on a case-by-case basis, it is unclear how this could be performed in a systematic and unbiased manner across a large number of motifs. This is especially challenging, because prior biological knowledge indicates that multiple biological processes, including completion of postnatal liver development, acute and sustained xenobiotic responses to PB treatment and tumor promotion,

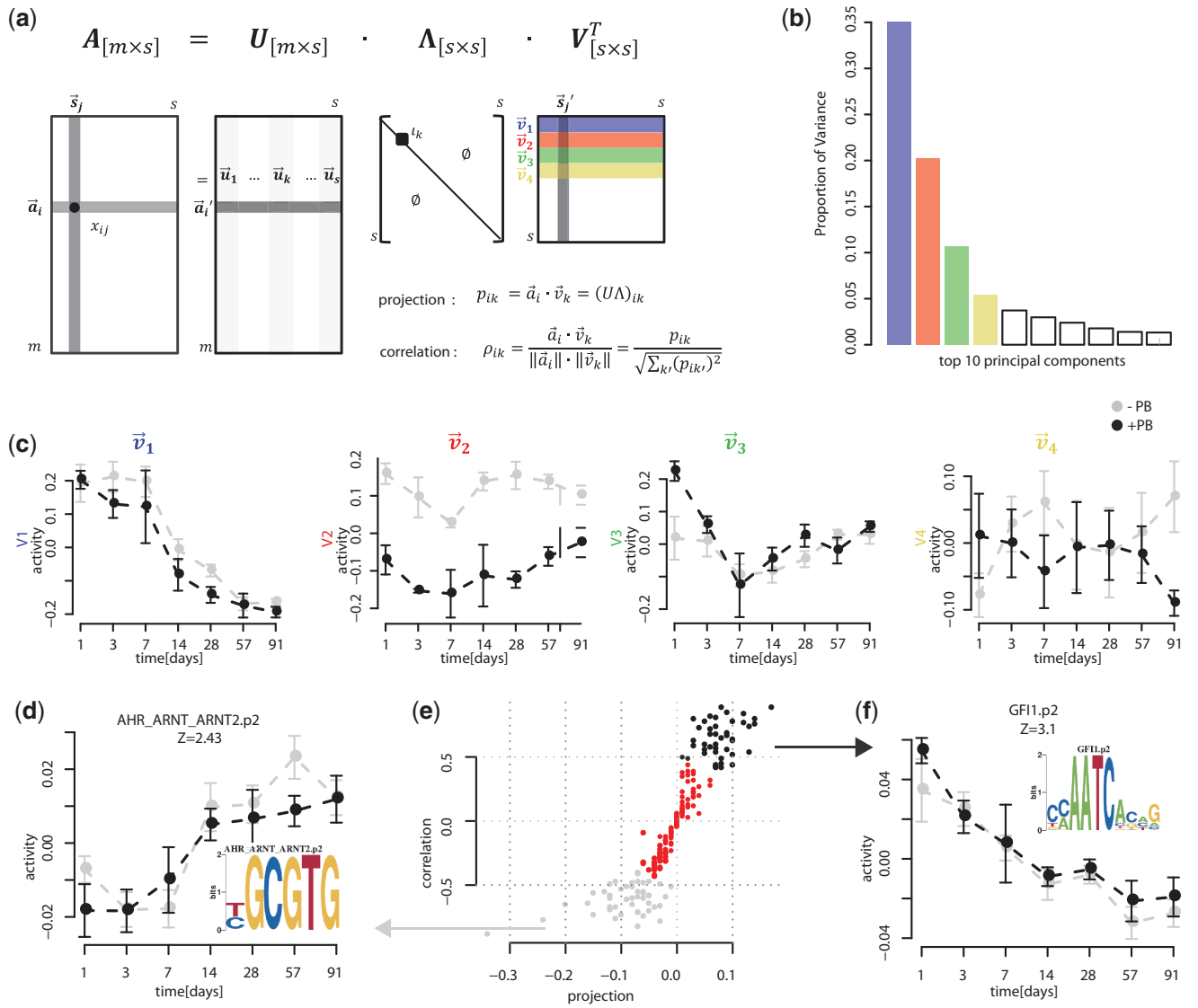

Figure 3. Overview of the analysis strategy for identifying key regulatory activities of the early PB-mediated transcriptional dynamics. (**a**) SVD factorizes the activity matrix of the early kinetic study: $\mathbf{A} = \mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{V}^T$, with the right singular vectors $\vec{v}_k$ giving orthonormal motif activity profiles that capture most of the variation in activity profiles across all motifs. (**b**) Proportion of the variance of the motif activity matrix explained by the 10 first components. The first (blue bar), second (red bar), third (green bar) and fourth (yellow bar) components account for 35, 20, 10 and 5%, respectively of the variance. (**c**) Activity profiles of the first four right singular vectors $\vec{v}_1$ through $\vec{v}_4$. Gray points indicate activities for the control samples and black points indicate activities for the PB-treated samples. (**d, f**) Examples of motif activity profiles that contribute and correlate negatively and positively, respectively, with the first right singular vector ($\vec{v}_1$). For each motif, a sequence logo representing its binding specificity is shown as an inset. (**e**) Scatter plot of the correlations ρ_{i1} and projections p_{i1} of all motifs i with the first right singular vector $\vec{v}_1$. Gray and black dots depict negatively and positively selected motifs.

are occurring in parallel in our system. To address this problem, we applied a SVD approach to decompose the matrix of inferred motif activities A_{ms} from the early kinetic study into linearly independent motif activity profiles that capture most of the variation in all motif activities.

Over 70% of the variance in the activity matrix was explained by the first four components of the SVD as evidenced by the spectrum of singular values (Figure 3b). The activity profiles of the first four right singular vectors, $\vec{v}_1$ through $\vec{v}_4$, are shown in Figure 3c. The first right singular vector accounted for 35% of the variance and was characterized by an approximately constant positive activity early in the time course that decreased dramatically after 2 weeks. The activity profile of this first singular vector was identical in the PB-treated and control groups. The steep drop in activity after 2 weeks coincided with the completion of postnatal liver development in this study, as indicated by the transcriptional profile of the hepatoblast marker α-fetoprotein (*Afp*) (47,48) (Supplementary Figure S4). We thus propose that this characteristic motif activity profile is associated with postnatal liver development. This conclusion is supported by some of the motifs associated with

this singular vector (see Supplementary Data). As this process is presumably not relevant for the process of non-genotoxic tumor promotion, we have not further focused on this singular vector and a characterization of its associated regulators is presented in the Supplementary Data.

The second singular vector accounted for 20% of the variance and was characterized by an activity profile that is almost entirely constant with time, but that showed a large difference between the PB-treated and vehicle-treated samples. This singular vector thus corresponds to a sustained xenobiotic response.

The third singular vector accounted for 10% of the variance and was characterized by a difference in activity between the control and treated group at Day 1 only; whereas activity in the control samples remained approximately constant in the first 3 days, activity was much higher at Day 1 and dropped significantly in PB-treated samples over the same initial phase. Given that PB mediates a transient mitotic response at Day 1 [also previously identified in other studies (18, 28)], we conclude that the biological pathway corresponding to this characteristic activity profile is the transient PB-mediated proliferative response.

Finally, the fourth singular vector accounted for 5% of the variance and was characterized by a divergence in the activity of the PB-treated and control groups in the last month of the 13-week time course. Given that this is the most significant singular value for differences between the PB-treated and control samples toward the end of the time course, we infer that this characteristic adaptive xenobiotic response activity profile might be an important contributor to the progressive creation of a tumor-prone environment.

In summary, we have shown that the behavior of regulatory motifs in the early stages of PB treatment are dominated by four characteristic activity profiles that account for >70% of variance of the motif activities and which correspond to the following fundamental biological processes: (i) the completion of postnatal liver development, (ii) a constant xenobiotic response, (iii) a PB-mediated acute mitogenic response and (iv) an adaptive xenobiotic response (late response to PB treatment).

Identification of representative motifs underlying the early dysregulated biological pathways

To determine motifs underlying the four characteristic motif activity profiles identified in the previous section, we selected motifs which contributed and correlated the most with each of the four singular vectors (Figure 3c, d, e and f). In this way we obtained, for each of the four singular vectors, two clusters of motifs with similar activity profiles, i.e. one correlating negatively with the singular vector and one correlating positively (Figure 3d and f). The advantage of extracting clusters of the most important regulatory motifs in this way, rather than simply clustering the motif activity profiles directly, is that many of the motif activity profiles contain components associated with different biological processes that are operating in parallel in our system. By first using SVD to identify the most significant characteristic activity profiles that are mutually 'independent', i.e. the singular vectors, we disentangle the regulatory activities associated with these different processes and cluster the motifs by the biological process.

We further refined the selection of the motifs associated with each singular vector as follows: (i) removing motifs for which the overall significance was too low ($z < 1.5$ for motifs regulating postnatal liver development or the constant xenobiotic response and under $z < 1.0$ for motifs regulating the transient mitogenic response or the adaptive xenobiotic response); (ii) removing motifs whose cognate TFs were not expressed in the liver (log expression <6.0); (iii) using z-scores for the differential activity per time point between PB-treated and control samples, we required $z_{\Delta c} \geq 1.5$ at minimum four time points out of the seven to belong to $\vec{v}_2$, at Day 1 to belong to $\vec{v}_3$ and at Day 91 to belong to $\vec{v}_4$. This lead to the identification of eight groups of motifs, i.e. two for each characteristic profile (Supplementary Table S1).

To further investigate the biological roles of the motifs associated with the four singular vectors, we performed Gene Ontology and KEGG functional enrichment analysis of the targets of the eight groups of motifs that are either positively or negatively associated with one of the singular vectors (Figure 4 and Supplementary Figure S5). Below is a brief description of most important findings.

Constant xenobiotic response. As discussed further below, it is well known that CAR is a crucial regulator involved in the xenobiotic response and thus a prime candidate for a regulator associated with a constant xenobiotic response. Unfortunately, as there is currently no high-quality regulatory motif available for CAR, our TFBS predictions do not include CAR target sites and our analysis is thus unable to infer CAR's activity *ab initio*. However, our analysis identified several additional regulators that are associated with a sustained xenobiotic response, i.e. a constant difference in activity between the PB-treated and control samples (a full list of associated motifs is presented in Supplementary Table S1).

Among these is TBP, whose targets are significantly upregulated under PB treatment and enriched in oxidation-reduction processes (Figure 4a), and NFE2 whose target genes are involved in homeostatic processes (Figure 4b) and include the proteasome complex (e.g. *Psmc3*, *Ufd1l* and *Ube2v1*) and oxidative stress genes (e.g. *Ggt1*, *Txn1* and *Adh7*). These targets represent key pathways of the liver drug-induced response that have been recently shown to be regulated by NFE2 in hepatocytes (49).

Transient proliferative response. It has been observed previously that PB treatment leads to a transient mitogenic response (18,28,20). Our analysis revealed that the process is positively regulated by the E2F family of TFs, whose motif activity is significantly increased at Day 1 upon PB treatment ($z = 2.2$). E2F family members are known regulators of cell proliferation and the functions of their predicted targets (Figure 4c) further confirms their specific role in DNA replication, DNA repair and mitosis (50–53).

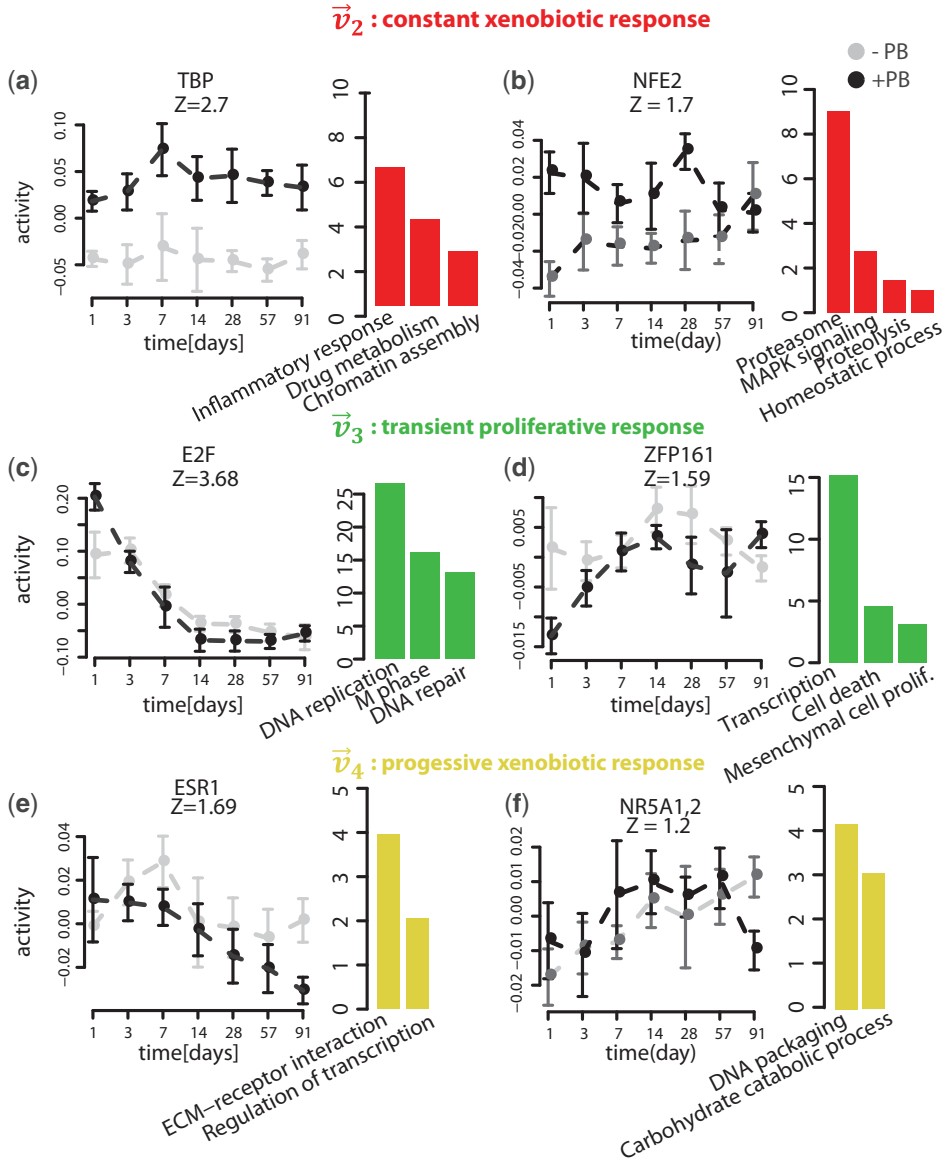

Figure 4. Examples of motif activity profiles that are associated with the constant xenobiotic response (**a,b**), the transient proliferative response (**c,d**) and the progressive xenobiotic response (**e,f**). For each motif, a *z*-value is indicated that quantifies the overall significance of the motif in the early kinetic dataset. In addition, for each motif a selection of biological pathways and functional categories (Gene Ontology or KEGG) that are enriched among target genes is plotted to the right of the activity profile. The size of each bar corresponds to the significance $[-\log_{10}(p-\text{value})]$ of the enrichment.

Interestingly, while eight TFs are potentially binding this motif, three of them (E2F1, E2F2 and E2F8) display a positive correlation between their gene expression and the motif activity in the time course (Supplementary Figure S3a and Supplementary Table S6), suggesting that it may be these three TFs that are involved in the PB-mediated transient hyperplastic response.

Our analysis predicted ZFP161 as an additional regulator of the transient hyperplastic response, whose targets are downregulated upon PB treatment. Interestingly, ZFP161's target genes are enriched in transcriptional repressors (e.g. *Rb1*, *Bcl6*, *Tle2*, *Klf9* and *Foxp1*), many of which are known to repress the cell cycle and cell growth. Moreover, positive regulation of cell proliferation by ZFP161 is further supported by negative regulation of cell death genes (Figure 4d). Together, these results suggest that PB-mediated ZFP161 activation may lead to the downregulation of an important group of transcriptional repressors and concomitant cell cycle activation.

Progressive xenobiotic response. Finally, our analysis identified several motifs associated with a divergence between motif activity in the PB-treated and control samples in the last month of the time course. Among the downregulated motifs is ESR1, whose predicted targets regulate extracellular matrix (ECM) genes and may thus regulate tissue remodeling (Figure 4e). NR5A1,2 is an additional regulator whose activity was downregulated

after 3 months of PB treatment (Figure 4e). Interestingly, Nr5a2 (known as liver receptor homolog-1 or LRH-1) is an established regulator of cholesterol, bile acid homeostasis, glucose and lipid metabolism (54,55), as confirmed by predicted targets functions in carbohydrate metabolism (Figure 4f).

Regulators of PB-mediated long-term liver gene expression changes are downstream of CAR signaling

In order to assess the importance of CAR in the livers in physiological conditions, i.e. without PB treatment and to identify to what extent the response to PB treatment is downstream of CAR activation, we made use of gene expression profiles from CAR WT and KO mice (19). We first identified regulators that are downstream of CAR under physiological conditions by comparing motif activities between non-treated CAR KO and WT samples (Figure 2a provides a schematic representation of all motif activity contrasts that we calculated). Only five motifs were significantly downregulated in their activity upon CAR deletion (Supplementary Table S2 provides a full list). To assess the CAR dependence of the regulatory motif changes mediated by PB treatment, we compared regulatory motifs that are perturbed in activity upon PB treatment in WT animals, with motifs that are perturbed upon PB treatment in CAR KO animals. Strikingly, of the 23 motifs dysregulated upon PB treatment in WT mice, none was dysregulated in KO mice, indicating that all regulators of PB-mediated gene expression changes at Day 161 are downstream of CAR signaling (Supplementary Table S2). This result is in line with previous studies where CAR was shown to be critical for both the acute (20) and chronic (19) transcriptional response to PB treatment. These results are further confirmed by an SVD analysis (Supplementary Results and Supplementary Figure S6), which shows that the major source of motif activity changes in these liver samples is the CAR-dependent liver response to PB treatment. In summary, it is highly likely that all motif activity changes observed in the early response time course also depend on CAR activation.

Identification of specific regulators of promoted tumors involved in early PB-mediated response

Our analysis above has focused on regulators that are perturbed during the first 3 months of PB treatment, whereas it takes several more months for tumors to be detected at the histopathologic level (21). We next investigated which regulators have different activities in the end-stage tumors that are observed after 8 months of treatment, as compared with their surrounding tissue (Figure 2b). We hypothesized that motifs perturbed both in the early response as well in the end-stage tumors may likely be involved in the process of tumor formation. Moreover, we distinguished 'promoted' tumors, which are characterized by mutations that cause constitutive activation of β-catenin, from 'non-promoted' tumors that are characterized by mutations in Ha-ras activation. Motifs that are perturbed in promoted tumors, but not in Ha-ras tumors, are prime candidates for involvement in the non-genotoxic tumor promotion.

We find eight motifs that are perturbed in both promoted and non-promoted tumors (Supplementary Table S3). Half of these were also associated with one of the singular vectors of the early PB treatment time course. In particular, the motif NR5A1,2 was associated with singular vector 4, showing a downregulation in the PB-treated animals in the third month of the time course, is also downregulated in the end-stage tumors. Predicted targets for NR5A1,2 are involved in several known metabolic functions of the liver [oxido-reduction processes, peroxisome proliferator-activated receptor (PPAR) signaling, and energy metabolism] (55), consistent with target functions at the early time points, indicating that NR5A1,2 downregulation is associated with hepatocyte loss of function (Figure 5a). Furthermore, our analysis identifies SOX{8,9,10} as a regulator of cell proliferation in both promoted and non-promoted tumors (Figure 5a). As these motifs are perturbed in both promoted and non-promoted tumors, they likely regulate genes involved in general liver tumor biology and are presumably not relevant for the specific process of non-genotoxic tumor promotion.

Seven motifs were dysregulated in promoted tumors only (Supplementary Table S4). Strikingly, all but one of these motifs were associated with one of the singular vectors of the early PB treatment time course. In particular, the E2F motif, which we found to be a positive regulator of both the postnatal liver growth and the transient PB-mediated mitogenic response, is here observed to be upregulated only in promoted tumors ($z_{\text{promoted tum.}} = 2.6$), showing no significant perturbation in the non-promoted tumors ($z_{\text{non-promoted tum.}} = -0.3$). Importantly, E2F1, E2F2 and E2F8, previously identified as strong candidate regulators of early PB-mediated transient hyperplastic response, display similar positive correlation between their gene expression and the motif activity in the tumor (Supplementary Figure S3a and Supplementary Table S5). Notably, the cellular functions regulated by SOX{8,9,10} and E2F at tumor stage (Figure 5a and b) suggests that these motifs have distinct regulatory effects on cell proliferation; while SOX{8,9,10} regulates mitosis, E2F targets specifically regulate DNA replication (Supplementary Figure S2).

The ZFP161 motif, which we found to negatively regulate transcriptional repressors of the cell cycle in the early stages of PB treatment, also displays significant decrease in activity in promoted tumors ($z_{\text{promoted tum.}} = -1.6$), but not in non-promoted tumors ($z_{\text{non-promoted tum.}} = 0.7$) (Figure 5b). Interestingly, these results suggest that similar regulatory mechanisms, involving E2F and ZPF161, are responsible for the proliferation that occurs transiently immediately upon PB treatment as well as the proliferation in promoted tumors. Moreover, this upregulation of proliferation, which might involve the release of specific cell-cycle checkpoints, is clearly distinct from the regulatory mechanism responsible for upregulation of proliferation in the non-promoted tumors.

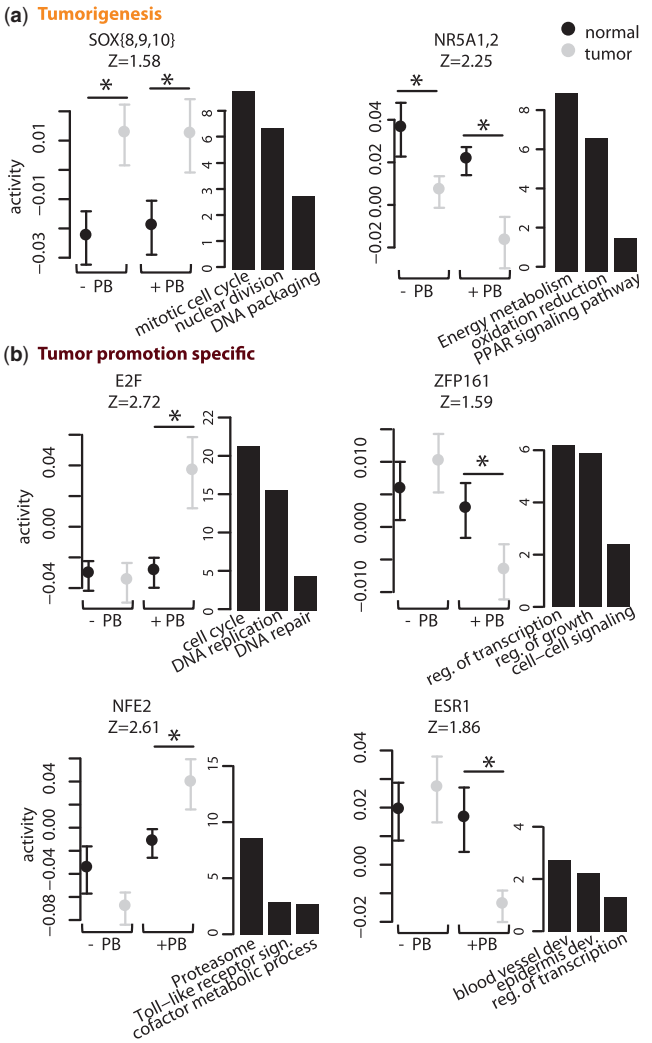

Figure 5. Regulators of liver tumorigenesis and tumor promotion. (**a**) Activities of two regulators that are dysregulated in both promoted and non-promoted tumors. (**b**) Activities of four regulators that are specifically dysregulated in promoted tumors. For each regulator, the activities in the tumor and surrounding normal tissue are indicated by black and turquoise points, respectively. A z-value quantifying the overall significance of the motif in tumor dataset is indicated below each motif's name. A selection of biological pathways and functional categories (Gene Ontology or KEGG) enriched among target genes of these motifs are shown on the right of each activity profile. The height of each bar corresponds to the significance $[-\log_{10}(P-\text{value})]$ of the enrichment. Differences in activity between the tumor and surrounding tissues that are significant are indicated by an asterisk ($|z_{\Delta act.}| \geq 1.5$).

Another motif specifically upregulated in promoted tumors is NFE2 ($z_{\text{promoted tum.}} = 2.5$ versus $z_{\text{non-promoted tum.}} = -1.4$). Furthermore targets of NFE2 that already showed upregulation at the early stage, such as several members of the protease family and oxidative stress response, e.g. *Aox1*, *Acox2*, *Srxn1* and *Mocos*, show continued activation in the promoted tumors (Figure 5b).

Finally, our analysis revealed a significant decrease in activity of the motif bound by ESR1 in promoted tumors only ($z_{\text{promoted tum.}} = -2.9$ versus $z_{\text{non-promoted tum.}} = 0.5$). Moreover, our analysis shows ESR1 regulation of genes involved in anatomical structure morphogenesis/tissue remodeling (genes, e.g. coding for collagen and fibronectin) that are progressively downregulated upon PB treatment and remain repressed at the tumor stage (Figure 5b). We also performed an SVD analysis of the activity matrix of this dataset (Supplementary Results and Supplementary Figure S7). The analysis identified the most significant singular component with regulators of promoted tumors that largely overlap those identified by differential motif activity analysis. The second singular component identified a number of regulators of liver tumorigenesis. Interestingly, these motifs were not identified by differential motif activity analysis, suggesting that SVD analysis can identify a significant effect of a set of motifs even when the differential activity of each motif is not significant by itself.

Early regulators of liver tumor promotion downstream of β-catenin signaling

It has been established that liver tumor promotion by PB requires functional β-catenin (56) and promoted tumors are characterized by mutations that cause constitutive activation of β-catenin. However, it remains unclear how PB promotes the outgrowth of pre-existing β-catenin activated cells. The ability for β-catenin to physically interact with various co-factors and nuclear receptors (57,58) suggests that the predicted regulators of PB-mediated liver tumor promotion may interact with the β-catenin pathway.

We thus investigated which regulators are downstream of β-catenin under physiological conditions by comparing motif activities in non-treated WT and β-catenin KO cells (Figure 2c). This analysis showed massive changes in regulatory activities upon KO of β-catenin, with as many as 33 motifs significantly perturbed in their activity (Supplementary Table S5). Note that this analysis successfully retrieved known co-factors of β-catenin such as the Tcf7-Lef1 motif, whose activity decreases strongly upon β-catenin KO ($z_{\text{ko-wt}} = -3.2$). Furthermore, two of the previously identified regulators of liver tumor promotion, i.e. E2F ($z_{\text{ko-wt}} = -2.0$) and NFE2 ($z_{\text{ko-wt}} = -2.2$), were negatively modulated upon β-catenin KO, whereas ESR1 ($z_{\text{ko-wt}} = 2.9$) was positively modulated. These findings support the hypothesis of a positive interaction between E2F/NFE2 and the β-catenin signaling pathway. The strong positive correlation between ESR1 gene expression and motif activity in both this study and the tumor study (Supplementary Figure S3 and Supplementary Table S6) supports a negative interaction between the β-catenin signaling pathway and ESR1 in liver, potentially through direct repression of target gene by β-catenin.

DISCUSSION

Here we describe a novel bioinformatics approach for the automated identification of independent transcription regulatory programs within a complex *in vivo* tissue environment. Using well-characterized mouse mechanistic models for non-genotoxic hepatocarcinogeneis, we were able to successfully infer the contributions of key

regulators of phenobarbital-mediated xenobiotic responses, tumor promotion and end-stage tumors as well as assess their dependence on the CAR and β-catenin signaling pathways.

Motif activity response analysis, which models observed gene expression patterns in terms of computationally predicted TF-binding sites, has been specifically designed to identify the key regulators responsible for the observed gene expression dynamics. One of its strengths is that MARA does not rely directly on the mRNA expression of the TFs, but instead infers the activities of regulators from the expression of their predicted target genes. Consequently, MARA can easily identify changes in motif activities that are due to post-translational modifications, changes in cellular localization or interactions with co-factors. This is specifically relevant for our model system in which PB indirectly triggers changes in gene expression via EGFR signaling-mediated post-translational modification and nuclear translocation of the TF CAR (15,16).

A major challenge in the analysis of the complicated *in vivo* systems such as the one we study here, is that the observed genome-wide expression changes result from multiple biological pathways dynamically changing in parallel. Consequently, even when MARA allows us to infer the regulatory activities of key TFs across the samples, it may be challenging to identify the independent biological processes that these regulators contribute to and how each regulator is contributing to each process. To address this, we here developed a new analysis approach based on SVD that decomposes the entire matrix of motif activities across all samples and identifies the major mutually independent activity profiles.

Our results show that this approach successfully identifies the major biological pathways underlying the response to PB treatment and it furthermore allows us to identify how the key regulators are contributing to each of these pathways. We identified the roles of E2F and ZFP161 in the regulation of cell proliferation in both the early transient mitogenic response and specifically in promoted tumors. We identified ESR1 as a key regulator of establishing a tumor-prone environment and we identified NFE2 as a key regulator of the sustained xenobiotic response. Figure 6 schematically summarizes these key findings, showing both the overall picture that emerges of the biological processes involved in PB-mediated tumor promotion (Figure 6a) as well as the key regulators that we identified and their role in the various processes (Figure 6b).

In the next sections we discuss these key findings, put them into context of relevant available literature and put forward concrete hypotheses for the biological mechanisms involved in these regulatory processes. Finally, where possible, we also discuss pieces of supporting evidence for the hypotheses we put forward.

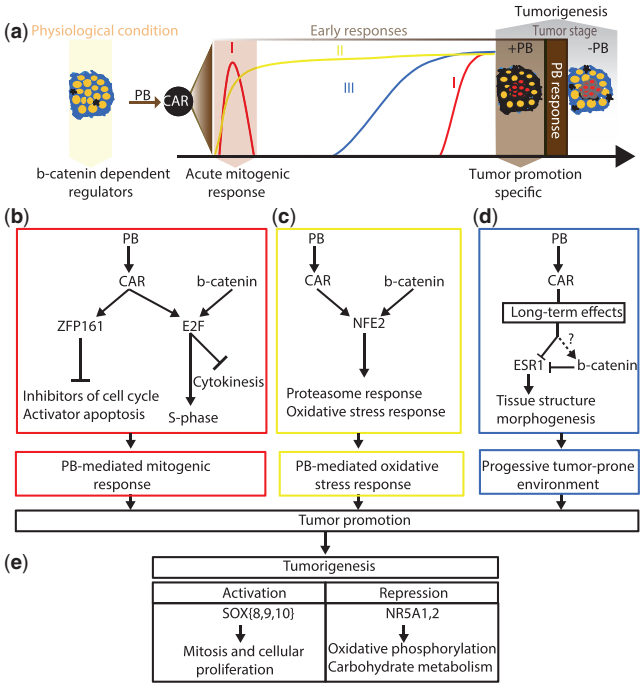

Figure 6. Schematical representation of PB-mediated tumor promotion, as it emerges from our study. (**a**) Illustration of the PB-mediated tumor promotion process and the aspects elucidated by the four experimental studies that we analyze. KO of β-catenin identifies regulators downstream of β-catenin in physiological conditions (yellow arrow). This study and previous analyses suggest that all regulatory effects of PB treatment are downstream of CAR activation (brown arrow and black circle). This study's motif activity and SVD analysis of the early kinetic time course identified three key biological processes induced by PB treatment: a transient mitogenic response, which is also associated with a late resurgence of proliferation (I, red), a sustained xenobiotic response (II, yellow) and a late response which is likely involved in establishing a tumor-prone environment (III, blue). Comparison of promoted and non-promoted tumors identifies motifs dysregulated in all tumors and in promoted tumors only (gray arrows). (**b–d**) Summary of the key regulators of liver tumor promotion organized according to biological processes (colored boxes matching the colors of processes I, II and III in panel a) with arrows indicating regulatory interactions between regulators and on selected target genes. (b) E2F and ZFP161 regulate PB-mediated hepatocyte proliferation at the early and promoted tumor stage. E2F is downstream of β-catenin signaling and likely induces both DNA replication, via upregulation of *E2f1,2* and aborted cytokinesis via upregulation of *E2f8* and *c-myc*. ZFP161 is likely involved in the G0–G1 transition via transcriptional repression of transcriptional repressors of cell growth and cell cycle. (c) NFE2, downstream of β-catenin as well is involved in the sustained xenobiotic response, upregulating proteasome activity and the oxidative stress response. (d) PB-mediated suppression of ESR1 activity underlies development of a tumor-prone environment, most likely through repression of tissue morphogenesis. β-Catenin signaling represses ESR1. (**e**) Key regulators involved in tumorigenesis, i.e. dysregulated in both promoted and non-promoted tumors. Increased SOX{8,9,10} activity likely regulates hepatocyte mitosis and proliferation via upregulation of cyclins. Decrease in NR5A1,2 activity is detected after 3 months of PB treatment and maintained in tumor samples and therefore a good early indicator of hepatocyte loss-of-function associated with tumorigenesis.

E2F as a positive regulator of the PB-mediated proliferative response at both the early and tumor stages

An important aspect of PB-mediated tumor promotion is the ability of PB to induce a transient mitogenic response and to cause liver neoplasia upon chronic administration. However, the exact mechanisms responsible for the exit from the quiescent state and the re-entry into the cell cycle remain largely unknown [see (59) for a review]. Our analysis revealed that the regulatory motif bound

by the E2F family of TFs is one of the key factors positively contributing to the early proliferative response upon PB treatment. In addition, E2F is upregulated in promoted tumors, but not in non-promoted tumors. Importantly, the absence of E2F motif modulation in non-promoted tumors argues against the hypothesis that the motif is simply reflecting increased proliferative activity. Furthermore, the fact that KO of β-catenin in physiological conditions leads to downregulation of E2F activity implies that β-catenin positively regulates E2F activity (either directly or indirectly) and suggests that PB-mediated activation of β-catenin may contribute to the upregulation of E2F activity at the tumor stage.

The plausibility of a role for E2F TFs in PB-mediated tumor promotion is supported by numerous studies reporting a central role of distinct E2F family members in hepatocellular carcinoma (60,61). More specifically, PB-mediated modulation of E2F gene regulation in freshly isolated hepatocytes has been previously suggested (62). Here we show a highly specific upregulation of E2F activity in promoted tumors and a potential role in tumor promotion through β-catenin-mediated activation.

The E2F family contains eight different TFs that can bind to the E2F motif and the MARA analysis does not directly predict which of these eight TFs is mainly responsible for the activity of the E2F regulatory motif in this system. However, measurements of motif activity correlation with mRNA expression of the TFs (Supplementary Figure S4 and Supplementary Table S6) shows that the expression of E2F1, E2F2 and E2F8 exhibit the most significant correlation with E2F motif activity in the time course and tumor studies. This makes these TFs the most likely candidates for driving the E2F motif activity, but it should be noted that motif activity changes do not necessarily require changes in mRNA levels of the binding TFs, i.e. the activity change may be due to post-translational modifications, nuclear localization, etc. E2F7 and E2F8 have been recently shown to play a key role in positively regulating hepatocyte polyploidy (63,64). Interestingly, Myc has been shown to be an additional positive regulator of polyploidy in hepatocytes (65,66). Furthermore, both *E2f8* and *c-myc* are significantly upregulated in promoted tumors only and both are predicted targets of E2F. Given that ligands of nuclear receptors such as PB and TCPOBOP have been shown to cause liver polyploidization (59,67,68), we propose that both E2F1 and E2F8 are responsible for the E2F activity modulation at the tumor stage and that they regulate distinct cell cycle checkpoints, in particular, regulation of entry in S-phase for E2F1 and inhibition of cytokinesis for E2F8 together with Myc (Figure 6b).

ZFP161 as transcriptional repressor involved in the PB-mediated proliferative response at both the early and tumor stages

Our analysis revealed a decrease in activity of the motif bound by ZFP161 (also known as ZF5), i.e. an overall downregulation of its predicted targets upon PB treatment contributing to the early transient proliferative response. In addition, ZFP161 targets are downregulated in promoted tumors, but not in non-promoted tumors. Affymetrix gene expression analysis shows that while ZFP161 is not transcriptionally regulated by PB and its mRNA expression is not correlated with motif activity (Supplementary Figure S3), it is clearly expressed in the liver ($\log_2 e \approx 8.0$).

Although ZFP161 has been shown to be preferentially active in differentiated tissues with little mitotic activity (69), where it was shown to act as a transcriptional repressor of *c-myc* (70,71), we here show an increase in ZFP161 transcriptional repression of target genes enriched in transcriptional repressors (i.e. *Mxi1* and *Klf10*), several of these being negative regulators of cell cycle and cell growth. Therefore, we hypothesize that ZFP161 participates in the PB-mediated regulation of quiescent hepatocyte G0–G1 transition at both the early and tumor stages, by repressing negative regulators of cell cycle and positive regulators of apoptosis (Figure 6b).

The progressive PB-mediated downregulation of ESR1 contributes to establishing a tumor-prone environment

PB-mediated tumorigenesis involves dynamic changes in tissue composition, and the adaptive response of the liver to chronic stress eventually leads to the establishment of a tumor-prone environment. The identification of key factors that contribute to this process could provide valuable insight into the development of PB-mediated tumorigenesis. Our analysis identified ESR1 as a factor progressively downregulated upon chronic PB exposure, starting in the third month of PB treatment. In addition, ESR1 activity is downregulated in promoted tumors, but not in non-promoted tumors. These two observations make ESR1 a strong candidate regulator for the process of establishing a tumor-prone environment. Furthermore, β-catenin KO in physiological conditions leads to upregulation of ESR1 activity, implying that β-catenin represses (directly or indirectly) ESR1. Further supporting this direct link between β-catenin and ESR1 repression is the fact that the highest correlations between ESR1 activity and mRNA expression levels are observed in the β-catenin KO and tumor studies, i.e. precisely those experiments where β-catenin activity is predicted to change (Supplementary Figure S3). Importantly, a physical interaction between β-catenin/TCF-4 and ESR1 has already been reported in other physiological contexts (72,73). That ESR1 can have tumor suppressor activity is supported by various studies (4,74–77). However, here we propose more specifically that the progressive suppression of ESR1 activity from early hyperplastic tissue to cancer (78) is mediated by PB chronic exposure and is one of the mechanisms underlying PB-mediated liver tumor promotion due to negative regulation of tissue morphogenesis (Figure 6d).

NFE2 as a regulator of exacerbated xenobiotic response associated with promoted tumors

Our analysis revealed that PB treatment causes a constant upregulation of homeostatic processes via NFE2 activation of proteasome and oxidative stress biological processes during the early phases of treatment and that

this upregulation persists into promoted tumors (Figure 6c). Of note, NFE2 regulatory activity in homeostatic processes has been shown in a recent study (49). This upregulation of NFE2 in tumors compared to the surrounding tissue is specific to promoted tumors. Furthermore, the fact that NFE2 activity is downregulated upon β-catenin KO in physiological conditions strongly suggests that β-catenin signaling is positively regulating NFE2 activity. As β-catenin is also involved in the regulation of drug metabolizing enzymes in the liver (79–82), we hypothesize that NFE2 and β-catenin cooperate in regulating genes involved in drug metabolism and that the xenobiotic response is partly exacerbated in promoted samples upon constitutive activation of β-catenin, resulting in further upregulation of NFE2.

Regulators of liver tumorigenesis

Our analysis identified several regulators of liver tumorigenesis (Supplementary Table S3). Interestingly, NR5A1,2 downregulation is observed early in the process of tumor promotion (after 3 months of PB treatment). Given its apparent role in hepatocyte liver function regulation (Supplementary Figure S1) we hypothesize that NR5A1,2 is associated with hepatocyte loss of function (Figure 6e). SOX{8,9,10} is an additional regulator of liver tumorigenesis and our analysis indicates a role in hepatocyte proliferation. Finally, comparing functional enrichment between the target genes of SOX{8,9,10} and E2F at tumor stage revealed that while E2F specifically regulates DNA replication (Supplementary Figure S2), SOX{8,9,10} preferentially targets mitotic genes (Supplementary Figure S1). These results support our hypothesis that E2F targets cell cycle check points that are distinct from those shared with other tumors.

Future extensions of the modeling approach

In future work we will aim to address several limitations of the current modeling approach. First and foremost, the method is currently limited to inferring the activities of only those TFs for which sequence specificities are known, i.e. roughly 350 of the approximately 1500 mouse TFs. For example, we were not able to predict CAR motif activity as there is, to our knowledge, no high quality sequence motif available for CAR. This is not an intrinsic limitation of the method and as regulatory motifs for an increasing number of TFs becomes available, they can easily be incorporated into the method.

Another major limitation of MARA is that it currently focuses solely on predicted TFBSs in proximal promoters, ignoring the effects of distal enhancers. Although a number of combined experimental and computational methods have been put forward recently that allow genome-wide mapping of active enhancers [e.g. (83)], these methods require considerable investment and enhancer maps are only available for a small set of selected model systems. As the locations of relevant enhancers vary highly across tissues and model systems, successful incorporation of enhancers into MARA requires the availability of enhancer maps for the specific system under study.

Most importantly, all the hypotheses discussed in this work are based on analysis of high-througput data and future experimental studies will be required to characterize our inferred TF activities in more detail at the biochemical level. Such studies may include chromatin immunoprecipitation (ChIP) assays on liver tissue from control and phenobarbital-treated mice.

CONCLUSION

We have demonstrated that by combining motif activity response analysis with SVD, we are able to automatically untangle the regulatory activities underlying the perturbation of multiple biological pathways in complex *in vivo* systems and derive novel hypotheses regarding the key regulators and their role in the process. Our analyses provide novel mechanistic insight for PB-mediated tumor promotion in the mouse liver, including the identification of E2F and ZFP161 as regulators of PB-mediated hepatocyte proliferation at both early and tumor stages and progressive PB-mediated suppression of ESR1 activity that may contribute to the development of a tumor-prone environment. These findings may also help identify novel biomarkers for assessing the carcinogenic potential of xenobiotics.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online, Supplementary Material and Methods, Supplementary Results, Supplementary Figures S1–S7, Supplementary Tables S1–S7 and Supplementary References 1–20.

ACKNOWLEDGEMENTS

We thank Arne Müller, Florian Hahne, Michael Stadler, Lukas Burger and Audrey Kauffmann for bioinformatics support; Piotr Balwierz and Michail Pakov for input on MARA and MotEvo usage and adaptation of implementation; Ute Metzer, Markus Templin, Olivier Grenet, Harri Lempiaïnen for inputs in NGC model.

Authour contributions: Bioinformatic and statistical analyses were conducted by R.L. The manuscript was prepared by R.L., E.v.N., R.T, J.M., M.S. and J.G.

FUNDING

Innovative Medicine Initiative Joint Undertaking (IMI JU) (115001) (MARCAR project; http://www.imi-marcar.eu/); Novartis. EvN was supported by the Swiss Systems Biology Initiative SystemsX.ch within the network "Cellplasticity". Funding for open access charge: University of Basel.

Conflict of interest statement. None declared.

REFERENCES

1. Calvisi,D.F., Ladu,S., Gorden,A., Farina,M., Conner,E.A., Lee,J.S., Factor,V.M. and Thorgeirsson,S.S. (2006) Ubiquitous activation of Ras and Jak/Stat pathways in human HCC. *Gastroenterology*, **130**, 1117–1128.
2. Calvisi,D.F., Pinna,F., Ladu,S., Pellegrino,R., Simile,M.M., Frau,M., De Miglio,M.R., Tomasi,M.L., Sanna,V., Muroni,M.R. *et al.* (2009) Forkhead box M1B is a determinant of rat susceptibility to hepatocarcinogenesis and sustains ERK activity in human HCC. *Gut*, **58**, 679–687.
3. Malz,M., Weber,A., Singer,S., Riehmer,V., Bissinger,M., Riener,M.O., Longerich,T., Soll,C., Vogel,A., Angel,P. *et al.* (2009) Overexpression of far upstream element binding proteins: a mechanism regulating proliferation and migration in liver cancer cells. *Hepatology*, **50**, 1130–9.
4. Li,Z., Tuteja,G., Schug,J. and Kaestner,K.H. (2012) Foxa1 and Foxa2 are essential for sexual dimorphism in liver cancer. *Cell*, **148**, 72–83.
5. Suva,M.L., Riggi,N. and Bernstein,B.E. (2013) Epigenetic reprogramming in cancer. *Science*, **339**, 1567–1570.
6. Silva Lima,B. and Van der Laan,J.W. (2000) Mechanisms of nongenotoxic carcinogenesis and assessment of the human hazard. *Regul. Toxicol. Pharmacol.*, **32**, 135–143.
7. LeBaron,M.J., Rasoulpour,R.J., Klapacz,J., Ellis-Hutchings,R.G., Hollnagel,H.M. and Gollapudi,B.B. (2010) Epigenetics and chemical safety assessment. *Mutat. Res.*, **705**, 83–95.
8. Phillips,J.M., Yamamoto,Y., Negishi,M., Maronpot,R.R. and Goodman,J.I. (2007) Orphan nuclear receptor constitutive active/androstane receptor-mediated alterations in DNA methylation during phenobarbital promotion of liver tumorigenesis. *Toxicol. Sci.*, **96**, 72–82.
9. Thomson,J.P., Hunter,J.M., Lempiäinen,H., Müller,A., Terranova,R., Moggs,J.G. and Meehan,R.R. (2013) Dynamic changes in 5-hydroxymethylation signatures underpin early and late events in drug exposed liver. *Nucleic Acids Res.*, **41**, 5639–5654.
10. Cunningham,M.L. (1996) Role of increased DNA replication in the carcinogenic risk of nonmutagenic chemical carcinogens. *Mutat. Res.*, **365**, 59–69.
11. Williams,G.M. (2008) Application of mode-of-action considerations in human cancer risk assessment. *Toxicol. Lett.*, **180**, 75–80.
12. Lee,G.H. (2000) Paradoxical effects of phenobarbital on mouse hepatocarcinogenesis. *Toxicol. Pathol.*, **28**, 215–225.
13. Peraino,C., Fry,R.J. and Staffeldt,E. (1971) Reduction and enhancement by phenobarbital of hepatocarcinogenesis induced in the rat by 2-acetylaminofluorene. *Cancer Res.*, **31**, 1506–1512.
14. Peraino,C., Fry,R.J. and Staffeldt,E. (1973) Brief communication: Enhancement of spontaneous hepatic tumorigenesis in C3H mice by dietary phenobarbital. *J. Natl Cancer Inst.*, **51**, 1349–1350.
15. Kawamoto,T., Sueyoshi,T., Zelko,I., Moore,R., Washburn,K. and Negishi,M. (1999) Phenobarbital-responsive nuclear translocation of the receptor CAR in induction of the CYP2B gene. *Mol. Cell Biol.*, **19**, 6318–6322.
16. Mutoh,S., Sobhany,M., Moore,R., Perera,L., Pedersen,L., Sueyoshi,T. and Negishi,M. (2013) Phenobarbital indirectly activates the constitutive active androstane receptor (car) by inhibition of epidermal growth factor receptor signaling. *Sci. Signal*, **6**, ra31.
17. Diwan,B.A., Lubet,R.A., Ward,J.M., Hrabie,J.A. and Rice,J.M. (1992) Tumor-promoting and hepatocarcinogenic effects of 1,4-bis[2-(3,5-dichloropyridyloxy)] benzene (TCPOBOP) in DBA/2NCr and C57BL/6NCr mice and an apparent promoting effect on nasal cavity tumors but not on hepatocellular tumors in F344/NCr rats initiated with N-nitrosodiethylamine. *Carcinogenesis*, **13**, 1893–1901.
18. Huang,W., Zhang,J., Washington,M., Liu,J., Parant,J.M., Lozano,G. and Moore,D.D. (2005) Xenobiotic stress induces hepatomegaly and liver tumors via the nuclear receptor constitutive androstane receptor. *Mol. Endocrinol.*, **19**, 1646–1653.
19. Phillips,J.M., Burgoon,L.D. and Goodman,J.I. (2009) The constitutive active/androstane receptor facilitates unique phenobarbital-induced expression changes of genes involved in key pathways in precancerous liver and liver tumors. *Toxicol. Sci.*, **110**, 319–333.
20. Ross,J., Plummer,S.M., Rode,A., Scheer,N., Bower,C.C., Vogel,O., Henderson,C.J., Wolf,C.R. and Elcombe,C.R. (2010) Human constitutive androstane receptor (CAR) and pregnane X receptor (PXR) support the hypertrophic but not the hyperplastic response to the murine nongenotoxic hepatocarcinogens phenobarbital and chlordane in vivo. *Toxicol. Sci.*, **116**, 452–466.
21. Yamamoto,Y., Moore,R., Goldsworthy,T.L., Negishi,M. and Maronpot,R.R. (2004) The orphan nuclear receptor constitutive active/androstane receptor is essential for liver tumor promotion by phenobarbital in mice. *Cancer Res.*, **64**, 7197–7200.
22. Aydinlik,H., Nguyen,T.D., Moennikes,O., Buchmann,A. and Schwarz,M. (2001) Selective pressure during tumor promotion by phenobarbital leads to clonal outgrowth of beta-catenin-mutated mouse liver tumors. *Oncogene*, **20**, 7812–7816.
23. Palacios,J. and Gamallo,C. (1998) Mutations in the beta-catenin gene (CTNNB1) in endometrioid ovarian carcinomas. *Cancer Res.*, **58**, 1344–1347.
24. Morin,P.J. (1999) Beta-catenin signaling and cancer. *BioEssays*, **21**, 1021–1030.
25. Garcia-Rostan,G., Tallini,G., Herrero,A., D'Aquila,T.G., Carcangiu,M.L. and Rimm,D.L. (1999) Frequent mutation and nuclear localization of beta-catenin in anaplastic thyroid carcinoma. *Cancer Res.*, **59**, 1811–1815.
26. Garcia-Rostan,G., Camp,R.L., Herrero,A., Carcangiu,M.L., Rimm,D.L. and Tallini,G. (2001) Beta-catenin dysregulation in thyroid neoplasms: down-regulation, aberrant nuclear expression, and CTNNB1 exon 3 mutations are markers for aggressive tumor phenotypes and poor prognosis. *Am. J. Pathol.*, **158**, 987–96.
27. Tissier,F., Cavard,C., Groussin,L., Perlemoine,K., Fumey,G., Hagnere,A.M., Rene-Corail,F., Jullian,E., Gicquel,C., Bertagna,X. *et al.* (2005) Mutations of beta-catenin in adrenocortical tumors: activation of the Wnt signaling pathway is a frequent event in both benign and malignant adrenocortical tumors. *Cancer Res.*, **65**, 7622–7627.
28. Lempiainen,H., Couttet,P., Bolognani,F., Muller,A., Dubost,V., Luisier,R., Del Rio Espinola,A., Vitry,V., Unterberger,E.B., Thomson,J.B. *et al.* (2013) Identification of Dlk1-Dio3 imprinted gene cluster noncoding RNAs as novel candidate biomarkers for liver tumor promotion. *Toxicol. Sci.*, **131**, 375–386.
29. FANTOM Consortium, Suzuki, H., Forrest,A.R., van Nimwegen,E., Daub,C.O., Balwierz,P.J., Irvine,K.M., Lassman,T., Ravasi,T., Hasegawa,Y.*et al.* (2009) The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. *Nat. Genet.*, **41**, 553–562.
30. Arnold,P., Erb,I., Pachkov,M., Molina,N. and van Nimwegen,E. (2012) MotEvo: integrated Bayesian probabilistic methods for inferring regulatory sites and motifs on multiple alignments of DNA sequences. *Bioinformatics*, **28**, 487–494.
31. Pachkov,M., Balwierz,P.J., Arnold,P., Ozonov,E. and van Nimwegen,E. (2013) SwissRegulon, a database of genome-wide annotations of regulatory sites: recent updates. *Nucleic Acids Res.*, **41**, D214–D220.
32. Summers,K.M., Raza,S., van Nimwegen,E., Freeman,T.C. and Hume,D.A. (2010) Co-expression of FBN1 with mesenchyme-specific genes in mouse cell lines: implications for phenotypic variability in Marfan syndrome. *Eur. J. Hum. Genet.*, **18**, 1209–1215.
33. Aceto,N., Sausgruber,N., Brinkhaus,H., Gaidatzis,D., Martiny-Baron,G., Mazzarol,G., Confalonieri,S., Quarto,M., Hu,G., Balwierz,P.J. *et al.* (2012) Tyrosine phosphatase SHP2 promotes breast cancer progression and maintains tumor-initiating cells via activation of key transcription factors and a positive feedback signaling loop. *Nat. Med.*, **18**, 529–537.
34. Arner,E., Mejhert,N., Kulyté,A., Balwierz,P.J., Pachkov,M., Cormont,M., Lorente-Cebrián,S., Ehrlund,A., Laurencikiene,J., Hedén,P. *et al.* (2012) Adipose tissue microRNAs as regulators of CCL2 production in human obesity. *Diabetes*, **61**, 1986–1993.
35. Arnold,P., Scholer,A., Pachkov,M., Balwierz,P.J., Jorgensen,H., Stadler,M.B., van Nimwegen,E. and Schubeler,D. (2013) Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. *Genome Res.*, **23**, 60–73.

36. Perez-Schindler,J., Summermatter,S., Salatino,S., Zorzato,F., Beer,M., Balwierz,P.J., van Nimwegen,E., Feige,J.N., Auwerx,J. and Handschin,C. (2012) The corepressor NCoR1 antagonizes PGC-1a and estrogen-related receptor alpha in the regulation of skeletal muscle function and oxidative metabolism. *Mol. Cell. Biol.*, **32**, 4913–4924.
37. Tiwari,N., Meyer-Schaller,N., Arnold,P., Antoniadis,H., Pachkov,M., van Nimwegen,E. and Christofori,G. (2013) Klf4 Is a transcriptional regulator of genes critical for EMT, including Jnk1 (Mapk8). *PLoS One*, **8**, e57329.
38. Vervoort,S.J., Lourenco,A.R., van Boxtel,R. and Coffer,P.J. (2013) SOX4 mediates TGF-beta-induced expression of mesenchymal markers during mammary cell epithelial to mesenchymal transition. *PLoS One*, **8**, e53238.
39. Eisele,P.S., Salatino,S., Sobek,J., Hottiger,M.O. and Handschin,C. (2013) The peroxisome proliferator-activated receptor gamma coactivator 1alpha/beta (PGC-1) coactivators repress the transcriptional activity of NF-Kappab in skeletal muscle cells. *J. Biol. Chem.*, **288**, 6589.
40. Suzuki,T., Nakano-Ikegaya,M., Yabukami-Okuda,H., de Hoon,M., Severin,J., Saga-Hatano,S., Shin,J.W., Kubosaki,A., Simon,C., Hasegawa,Y. *et al.* (2012) Reconstruction of monocyte transcriptional regulatory network accompanies monocytic functions in human fibroblasts. *PLoS One*, **7**, e33474.
41. Hasegawa,R., Tomaru,Y., de Hoon,M., Suzuki,H., Hayashizaki,Y. and Shin,J.W. (2013) Identification of ZNF395 as a novel modulator of adipogenesis. *Exp. Cell Res.*, **319**, 68–76.
42. Meier-Abt,F., Milani,E., Roloff,T., Brinkhaus,H., Duss,S., Meyer,D.S., Klebba,I., Balwierz,P.J., van Nimwegen,E. and Bentires-Alj,M. (2013) Parity induces differentiation and reduces Wnt/Notch signaling ratio and proliferation potential of basal stem/progenitor cells isolated from mouse mammary epithelium. *Breast Cancer Res.*, **15**, R36.
43. Tiwari,N., Tiwari,V.K., Waldmeier,L., Balwierz,P.J., Arnold,P., Pachkov,M., Meyer-Schaller,N., Schubeler,D., van Nimwegen,E. and Christofori,G. (2013) Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming. *Cancer Cell*, **23**, 768–783.
44. R Core Team. (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
45. Huang,D.W., Sherman,B.T. and Lempicki,R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1–13.
46. Huang,D.W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
47. Camper,S.A. and Tilghman,S.M. (1989) Postnatal repression of the alpha-fetoprotein gene is enhancer independent. *Genes Dev.*, **3**, 537–546.
48. Tilghman,S.M. and Belayew,A. (1982) Transcriptional control of the murine albumin/alpha-fetoprotein locus during development. *Proc. Natl Acad. Sci. USA*, **79**, 5254–5257.
49. Lee,C.S., Ho,D.V. and Chan,J.Y. (2013) Nuclear factor-erythroid 2-related factor 1 regulates expression of proteasome genes in hepatocytes and protects against endoplasmic reticulum stress and steatosis in mice. *FEBS J.*, **280**, 3609–3620.
50. Helin,K. (1998) Regulation of cell proliferation by the E2F transcription factors. *Curr. Opin. Genet. Dev.*, **8**, 28–35.
51. Johnson,D.G., Schwarz,J.K., Cress,W.D. and Nevins,J.R. (1993) Expression of transcription factor E2F1 induces quiescent cells to enter S phase. *Nature*, **365**, 349–352.
52. Wu,L., Timmers,C., Maiti,B., Saavedra,H.I., Sang,L., Chong,G.T., Nuckolls,F., Giangrande,P., Wright,F.A., Field,S.J. *et al.* (2001) The E2F1-3 transcription factors are essential for cellular proliferation. *Nature*, **414**, 457–462.
53. Polager,S., Kalma,Y., Berkovich,E. and Ginsberg,D. (2002) E2Fs up-regulate expression of genes involved in DNA replication, DNA repair and mitosis. *Oncogene*, **21**, 437–446.
54. Mataki,C., Magnier,B.C., Houten,S.M., Annicotte,J.S., Argmann,C., Thomas,C., Overmars,H., Kulik,W., Metzger,D., Auwerx,J. *et al.* (2007) Compromised intestinal lipid absorption in mice with a liver-specific deficiency of liver receptor homolog 1. *Mol. Cell Biol.*, **27**, 8330–8339.
55. Oosterveer,M.H., Mataki,C., Yamamoto,H., Harach,T., Moullan,N., van Dijk,T.H., Ayuso,E., Bosch,F., Postic,C., Groen,A.K. *et al.* (2012) LRH-1-dependent glucose sensing determines intermediary metabolism in liver. *J. Clin. Invest.*, **122**, 2817–2826.
56. Rignall,B., Braeuning,A., Buchmann,A. and Schwarz,M. (2011) Tumor formation in liver of conditional beta-catenin-deficient mice exposed to a diethylnitrosamine/phenobarbital tumor promotion regimen. *Carcinogenesis*, **32**, 52–57.
57. Giera,S., Braeuning,A., Kohle,C., Bursch,W., Metzger,U., Buchmann,A. and Schwarz,M. (2010) Wnt/beta-catenin signaling activates and determines hepatic zonal expression of glutathione S-transferases in mouse liver. *Toxicol. Sci.*, **115**, 22–33.
58. Mulholland,D.J., Dedhar,S., Coetzee,G.A. and Nelson,C.C. (2005) Interaction of nuclear receptors with the Wnt/beta-catenin/Tcf signaling axis: Wnt you like to know? *Endocr. Rev*, **26**, 898–915.
59. Ledda-Columbano,G. and Columbano,A. (2011) Hepatocyte growth, proliferation and experimental carcinogenesis. In: Monga,S.P.S. (ed.), *Molecular Pathology of Liver Diseases,* vol. 5 of *Molecular Pathology Library*, Springer, USA, Chapter 54, pp. 791–813.
60. Conner,E.A., Lemmer,E.R., Omori,M., Wirth,P.J., Factor,V.M. and Thorgeirsson,S.S. (2000) Dual functions of E2F-1 in a transgenic mouse model of liver carcinogenesis. *Oncogene*, **19**, 5054–5062.
61. Palaiologou,M., Koskinas,J., Karanikolas,M., Fatourou,E. and Tiniakos,D.G. (2012) E2F-1 is overexpressed and pro-apoptotic in human hepatocellular carcinoma. *Virchows Arch.*, **460**, 439–446.
62. Gonzales,A.J., Christensen,J.G., Preston,R.J., Goldsworthy,T.L., Tlsty,T.D. and Fox,T.R. (1998) Attenuation of G1 checkpoint function by the non-genotoxic carcinogen phenobarbital. *Carcinogenesis*, **19**, 1173–1183.
63. Meserve,J.H. and Duronio,R.J. (2012) Atypical E2Fs drive atypical cell cycles. *Nat. Cell Biol.*, **14**, 1124–1125.
64. Pandit,S.K., Westendorp,B., Nantasanti,S., van Liere,E., Tooten,P.C.J., Cornelissen,P.W.A., Toussaint,M.J.M., Lamers,W.H. and de Bruin,A. (2012) E2F8 is essential for polyploidization in mammalian cells. *Nat. Cell Biol.*, **14**, 1181–1191.
65. Baena,E., Gandarillas,A., Vallespinos,M., Zanet,J., Bachs,O., Redondo,C., Fabregat,I., Martinez-A,C. and de Alboran,I.M. (2005) c-Myc regulates cell size and ploidy but is not essential for postnatal proliferation in liver. *Proc. Natl Acad. Sci. USA*, **102**, 7286–7291.
66. Conner,E.A., Lemmer,E.R., Sánchez,A., Factor,V.M. and Thorgeirsson,S.S. (2003) E2F1 blocks and c-Myc accelerates hepatic ploidy in transgenic mouse models. *Biochem. Biophys. Res. Commun.*, **302**, 114–120.
67. Marsman,D.S., Cattley,R.C., Conway,J.G. and Popp,J.A. (1988) Relationship of hepatic peroxisome proliferation and replicative DNA synthesis to the hepatocarcinogenicity of the peroxisome proliferators di(2-ethylhexyl)phthalate and 4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio.acetic acid (Wy-14,643) in rats. *Cancer Res.*, **48**, 6739–6744.
68. Melchiorri,C., Chieco,P., Zedda,A.I., Coni,P., Ledda-Columbano,G.M. and Columbano,A. (1993) Ploidy and nuclearity of rat hepatocytes after compensatory regeneration or mitogen-induced liver growth. *Carcinogenesis*, **14**, 1825–1830.
69. Numoto,M., Niwa,O., Kaplan,J., Wong,K.K., Merrell,K., Kamiya,K., Yanagihara,K. and Calame,K. (1993) Transcriptional repressor ZF5 identifies a new conserved domain in zinc finger proteins. *Nucleic Acids Res.*, **21**, 3767–3775.
70. Numoto,M., Yokoro,K., Yanagihara,K., Kamiya,K. and Niwa,O. (1995) Over-expressed ZF5 gene product, a c-myc-binding protein related to GL1-Kruppel protein, has a growth-suppressive activity in mouse cell lines. *Jpn J. Cancer Res.*, **86**, 277–283.
71. Sobek-Klocke,I., Disque-Kochem,C., Ronsiek,M., Klocke,R., Jockusch,H., Breuning,A., Ponstingl,H., Rojas,K., Overhauser,J. and Eichenlaub-Ritter,U. (1997) The human gene ZFP161 on 18p11.21-pter encodes a putative c-myc repressor and is homologous to murine Zfp161 (Chr 17) and Zfp161-rs1 (X Chr). *Genomics*, **43**, 156–164.

72. El-Tanani,M., Fernig,D.G., Barraclough,R., Green,C. and Rudland,P. (2001) Differential modulation of transcriptional activity of estrogen receptors by direct protein-protein interactions with the T cell factor family of transcription factors. *J. Biol. Chem.*, **276**, 41675–41682.
73. Kouzmenko,A.P., Takeyama,K., Ito,S., Furutani,T., Sawatsubashi,S., Maki,A., Suzuki,E., Kawasaki,Y., Akiyama,T., Tabata,T. *et al.* (2004) Wnt/beta-catenin and estrogen signaling converge in vivo. *J. Biol. Chem.*, **279**, 40255–40258.
74. Naugler,W.E., Sakurai,T., Kim,S., Maeda,S., Kim,K., Elsharkawy,A.M. and Karin,M. (2007) Gender disparity in liver cancer due to sex differences in MyD88-dependent IL-6 production. *Science*, **317**, 121–124.
75. Tsutsui,S., Yamamoto,R., Iishi,H., Tatsuta,M., Tsuji,M. and Terada,N. (1992) Promoting effect of ovariectomy on hepatocellular tumorigenesis induced in mice by 3'-methyl-4-dimethylaminoazobenzene. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.*, **62**, 371–375.
76. Shimizu,I., Yasuda,M., Mizobuchi,Y., Ma,Y.R., Liu,F., Shiba,M., Horie,T. and Ito,S. (1998) Suppressive effect of oestradiol on chemical hepatocarcinogenesis in rats. *Gut*, **42**, 112–119.
77. Yamamoto,R., Iishi,H., Tatsuta,M., Tsuji,M. and Terada,N. (1991) Roles of ovaries and testes in hepatocellular tumorigenesis induced in mice by 3'-methyl-4-dimethylaminoazobenzene. *Int. J. Cancer*, **49**, 83–88.
78. Eagon,P.K., Elm,M.S., Epley,M.J., Shinozuka,H. and Rao,K.N. (1996) Sex steroid metabolism and receptor status in hepatic hyperplasia and cancer in rats. *Gastroenterology*, **110**, 1199–1207.
79. Braeuning,A., Sanna,R., Huelsken,J. and Schwarz,M. (2009) Inducibility of drug-metabolizing enzymes by xenobiotics in mice with liver-specific knockout of Ctnnb1. *Drug Metab. Dispos.*, **37**, 1138–1145.
80. Chesire,D.R., Dunn,T.A., Ewing,C.M., Luo,J. and Isaacs,W.B. (2004) Identification of aryl hydrocarbon receptor as a putative Wnt/beta-catenin pathway target gene in prostate cancer cells. *Cancer Res.*, **64**, 2523–2533.
81. Colletti,M., Cicchini,C., Conigliaro,A., Santangelo,L., Alonzi,T., Pasquini,E., Tripodi,M. and Amicone,L. (2009) Convergence of Wnt signaling on the HNF4alpha-driven transcription in controlling liver zonation. *Gastroenterology*, **137**, 660–672.
82. Hailfinger,S., Jaworski,M., Braeuning,A., Buchmann,A. and Schwarz,M. (2006) Zonal gene expression in murine liver: lessons from tumors. *Hepatology*, **43**, 407–414.
83. Natarajan,A., Yardimci,G.G., Sheffield,N.C., Crawford,G.E. and Ohler,U. (2012) Predicting cell-type-specific gene expression from regions of open chromatin. *Genome Res.*, **22**, 1711–1722.

Supplementary Data

Raphaëlle Luisier [1], Elif B. Unterberger [2], Jay I. Goodman [3], Michael Schwarz [2], Jonathan Moggs [1], Rémi Terranova [1,*] and Erik Van Nimwegen [4,*]

[1]Discovery and Investigative Safety, Novartis Institutes for Biomedical Research, 4057 Basel, Switzerland

[2]Department of Toxicology, Institute of Experimental and Clinical Pharmacology and Toxicology, University of Tübingen, 72074 Tübingen, Germany

[3]Department of Pharmacology and Toxicology, Michigan State University, 48824 Michigan, U.S.A.,

[4]Biozentrum, University of Basel, and Swiss Institute of Bioinformatics, 4056 Basel, Switzerland

[*]To whom correspondence should be addressed: remi.terranova@novartis.com, erik.vannimwegen@unibas.ch

December 19, 2013

1 Supplementary Material and Methods

1.1 Gene expression datasets

A library of 109 genomewide mRNA expression patterns was compiled from four different studies (**Figure 1a**): 70 samples from a time series of expression data from liver samples of B6C3F1 vehicle- (i.e. control) or PB-treated mice at +1, +3, +7, +14, +28, +57 and +91 days of dosing (5 replicates) [1]; 8 mRNA expression patterns in livers of wild-type and hepatocyte-specific β-catenin knockout C3H/N [2] animals; 13 mRNA expression patterns in livers of wild-type and CAR knock-out C3H/N animals DEN-initiated at 5 weeks of age prior to 23 weeks of PB -or vehicle-treatment [3]. Datasets on global mRNA expression patterns (18 samples) from liver tumors and corresponding surrounding normal tissue of C3H/N animals DEN-initiated at 4 weeks of age prior to 35 weeks of PB- or vehicle-treatment were available to us from IMI-MARCAR partners (Unterberger *et al,(2013)*, manuscript submitted). Screening the tumors for mutations in Ha-ras, B-raf and Ctnnb1 (i.e. the β-catenin coding gene) confirmed that promoted tumors (from animals exposed to PB) were mutated in Ctnnb1 while non-promoted tumors were mutated in Ha-ras (data not shown, Unterberger *et al*, 2013). In all four studies gene expression was profiled using Affymetrix GeneChip MOE-4302 (Affymetrix, Santa Clara, CA) containing approximately 43,000 probe sets.

1.2 Affymetrix GeneChip processing

The analysis of the micro-array data was done with the R statistical package version 2.13 (2005) and Bioconductor libraries version 1.4.7 [4]. The four original data-sets containing Affymetrix CEL files were normalized independently using the Robust Multichip Average (RMA) implementation of the algorithm available in R/Bioconductor [4], producing four expression matrices, and the quality of the experiments was assessed using diverse statistics implemented in the package arrayQualityMetrics for R/Bioconductor [5].

2 Supplementary Results

2.1 Regulators associated with termination of developmental liver growth ($\vec{v}_1$)

To determine motifs underlying the four characteristic modes identified in this study, we selected motifs which contributed and correlated the most with each of the four singular vectors (**Figure 3c,d,e,f**). In this way we obtained, for each of the 4 singular vectors, two clusters of motifs with similar activity profiles, i.e. one correlating negatively with the singular vector, and one correlating positively (**Figures 3d,f**). We further refined the selection of the motifs associated with first singular vector as follows: 1) removing motifs for which the overall significance was lower than $z < 1.5$ and 2) removing motifs whose cognate TFs were not expressed in the liver (log-expression less than 6.0) $\log_2 e \leq 6.0$). This lead to the identification of 6 motifs motifs (**Supplementary Table S1**).

As originally observed in [1], completion of the post-natal liver development process occurs during the early PB-treatment time course, consisting in both hepatocyte proliferation at early stage, and progressive induction of liver-specific genes [6, 7]. We here identify key regulators of these two processes: 1. we show that post-natal liver growth (that decreases over time) is regulated by known regulators of cell proliferation such as the E2F family of TFs [8, 9, 10], SRF [11] and Myc [12, 13]; predicted target genes of these motifs have functions related to cell cycle and DNA replication (**Supplementary Figure S6i**), confirming the role of these regulators in cell proliferation. 2. We show that post-natal liver differentiation (which increases over time) is partly regulated by AHR, a known regulator of drug-metabolizing genes and transporters [14, 15, 16, 17] that has been shown to play key role in liver development [18]. Thus, the main biological process associated positively with the first singular vector is cellular proliferation associated with post-natal liver growth for the first two weeks of the time course. Conversely, the targets of the motifs that are negatively associated with the first singular vector, i.e. corresponding to genes that increase their expression after the first two weeks, are enriched for functions associated with hepatocyte terminal differentiation, such as 'liver development', 'drug metabolism' and 'transcriptional regulation'.

2.2 Singular value decomposition analysis of the activity matrix of the CAR KO data-set

In order to identify and quantify the sources of motif activity changes in the CAR KO data-set, we performed Singular Value Decomposition (SVD) of the activities of the 189 motifs across the four conditions (PB- and vehicle treated livers from wild-type and CAR KO mice). Over 50% of the variance in the activity matrix was explained by the first two components of the SVD as evidenced by the spectrum of singular values (**Figure S6a**).

In order to facilitate the biological interpretation of the singular vectors, we plotted the averaged activities of the right singular vectors v_{ks} over each of the four sample groups and further identified regulatory motifs whose activity profiles correlate most strongly (either positively or negatively) with the activity profile of the singular vector. Visualization of the averaged activity of the first two singular vectors $\vec{v}_1$ and $\vec{v}_2$ in each of the four sample groups is shown in **Figure S6b** and scatter plots of the correlations ρ_i and projections p_i of all motifs i with the first and second right singular vectors are shown in **Figure S6c**.

The first right singular vector accounts for 33% of the variance and is characterized by a positive activity upon PB treatment in wild-type animals only. Given the absence of positive activity in CAR KO treated animals, we propose that this component represents the liver response to PB that is CAR-dependent. Moreover motifs which contribute and correlate most strongly with the first singular vector (TBP, NFE2, REST, GLI1,2,3, FOSL2, ELK1,2, and ZNF143) are all down-stream of CAR signaling under PB treatment (**Table S2**) except CTCF, RXRG-dimer and STAT5{A,B}, further supporting the association of this component with the CAR-dependent liver response to PB treatment.

The second right singular vector accounts for 18% of the variance and is characterized by 1) a lower activity in wild-type liver samples compared to CAR KO samples, and 2) by an activity further lowered upon PB treatment in both wild-type and CAR KO samples (**Figure S6b**). We propose that this component represents the basal liver activity down-stream of CAR that is further exacerbated upon PB treatment. However the motifs that contribute and correlate most strongly with the second singular vector do not coincide with any of the 5 motifs identified by differential motif activity analysis as down-stream of CAR signaling under physiological condition (**Table S2**). Furthermore the average activities have large associated error-bars for each sample group, indicating that the interpretation of this component must be considered with caution.

In conclusion, the SVD-based analysis of the activity matrix of the CAR KO data-set indicates that the major

source of motif activity changes in these liver samples is the CAR-dependent liver response to PB treatment. This result is in line with the analysis based on differential motif activity. Importantly, prior biological knowledge indicates that at least two biological processes are occurring in this system, i.e the CAR KO effect and the xenobiotic response to PB treatment. Differential motif activity previously showed only a very minor CAR KO effect (only 5 motifs identified as down-stream of CAR signaling under physiological condition, see **Table S2**) that may explain the absence of strong association of any component with this biological process.

2.3 Singular value decomposition analysis of the activity matrix of the tumor study data-set

In order to identify and quantify the sources of motif activity changes in the tumor data-set, we performed Singular Value Decomposition (SVD) of the activities of the 189 motifs across the four conditions (PB- and vehicle treated normal and tumorigenic liver samples). Over 57% of the variance in the activity matrix was explained by the first two components of the SVD that are the two significant components of the matrix, as evidenced by the spectrum of singular values (**Figure S7a**).

In order to facilitate the biological interpretation of the singular vectors, we plotted the averaged activities of the right singular vectors v_{ks} over each of the four sample groups and further identified regulatory motifs whose activity profiles correlate most strongly (either positively or negatively) with the activity profile of the singular vector. Visualization of the averaged activity of the first two singular vectors $\vec{v}_1$ and $\vec{v}_2$ in each of the four sample groups is shown in **Figure S7b** and scatter plots of the correlations ρ_i and projections p_i of all motifs i with the first and second right singular vectors are shown in **Figure S7c**.

The first right singular vector accounts for 32% of the variance (**Figure S7a**) and is characterized by 1) a higher activity in PB-treated samples relative to non-treated samples, 2) an increased positive activity in promoted tumor samples relative to all other sample groups (normal treated and non-treated samples, and non-treated tumor samples) and 3) a slight decreased activity in non-promoted tumor samples relative to surrounding normal tissue (**Figure S7b**). Moreover several motifs which contribute and correlate most strongly with the first singular vector (NFE2, E2F1-5, PBX1, and ESR1) as depicted in **Figure S7c**, have been identified as specific regulators of promoted tumors by differential motif activity analysis (see **Table S4**). These results indicate that motifs associated with this component are generally associated with a response to PB treatment which is further 1) exacerbated in promoted tumor samples and 2) inhibited in non-treated tumor samples, suggesting that the first component captures motifs associated with biological pathways underlying promoted tumors that are already up-regulated upon PB treatment and down-regulated in non-promoted tumors.

The second right singular vector accounts for 25% of the variance (**Figure S7a**) and is characterized by an overall decreased activity in tumor samples relative to normal samples, irrespective of the PB treatment (**Figure S7b**); this suggests that the second component captures motifs associated with biological pathways underlying tumorigenesis. It is however noteworthy that none of the motifs which contribute and correlate most strongly with the second singular vector (**Figure S7c**) were identified as regulators of tumorigenesis by differential motif activity analysis (**Table S3**). One explanation for this could be a strong variability in activity profiles leading to low Z-value of differential activity.

In conclusion the SVD-based analysis of the activity matrix allows for the identification of 1) regulators of promoted tumors (first component) which are consistent with those identified by differential motif activity analysis, and 2) regulators of liver tumorigenesis, which were not identified by differential motif activity analysis, potentially due to high noise to signal ratio.

3 Supplementary figures

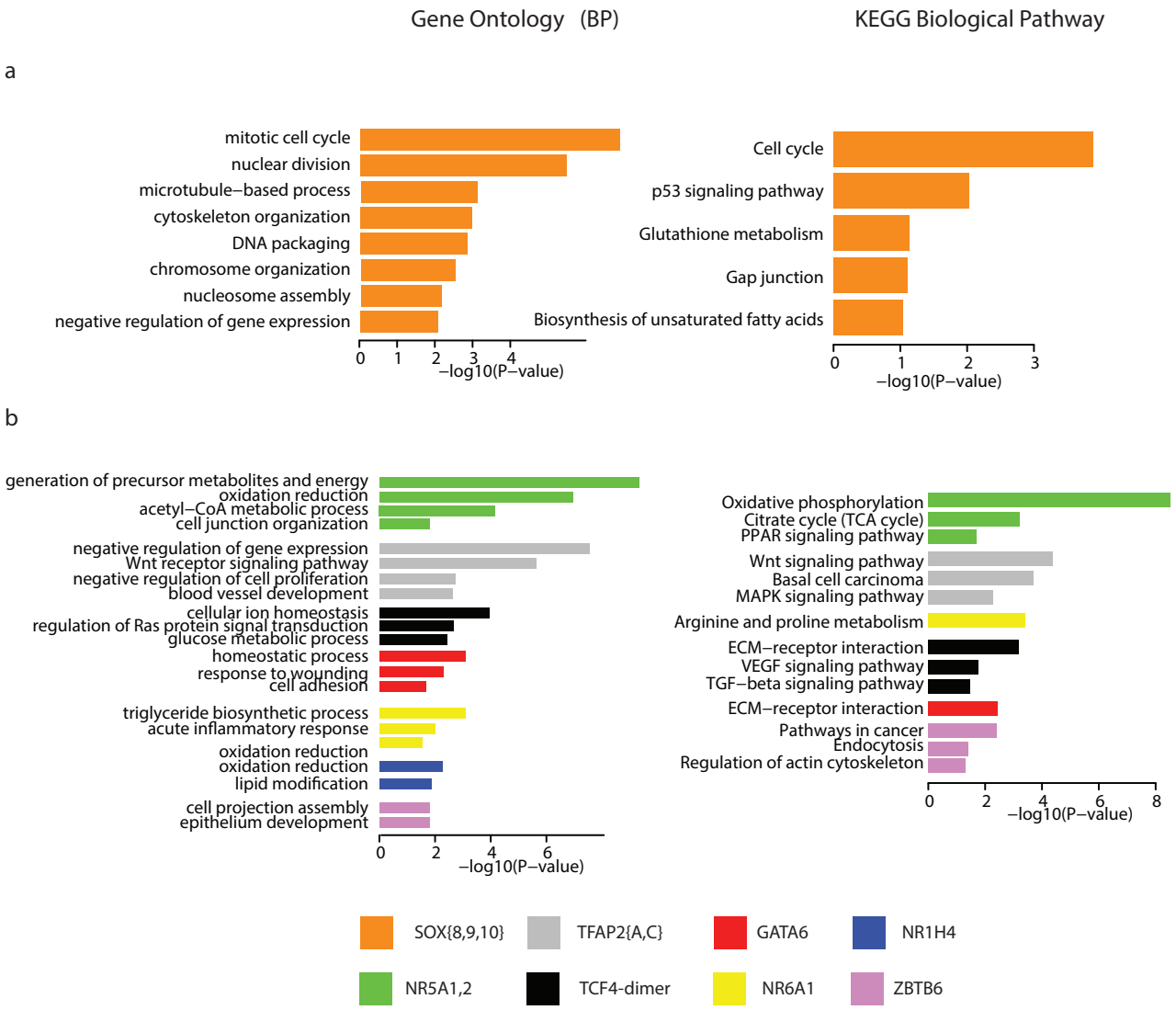

Figure S1: Selection of representative biological terms and processes associated with the predicted target genes of motifs which activities were significantly **(a)** higher or **(b)** lower in promoted tumors relative to surrounding treated normal tissue, and in non-promoted tumors relative to surrounding non-treated normal tissue (**Supplementary Table S3**). Bars are colored according to motif to which the target genes are associated with. Bar height indicates significance of functional enrichment as it represents the $-\log10(P-\text{Value})$ of functional enrichment in the given biological term or process as obtained from the DAVID Bioinformatic Resource (Database for Annotation, Visualization and Integrated Discovery) [19, 20] version 6.7, sponsored by the National Institute of Allergy and Infectious Disease (NIAID), NIH.

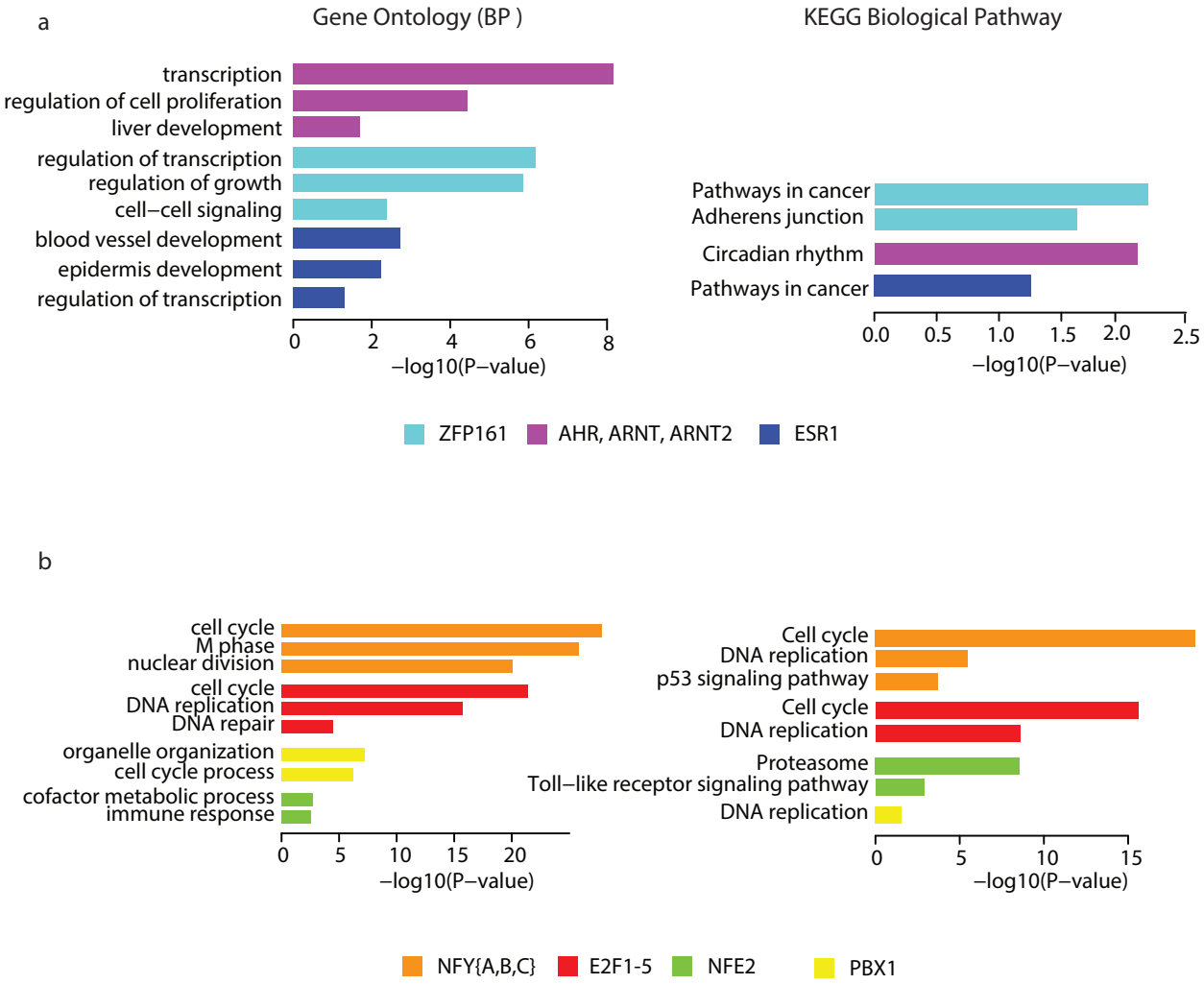

Figure S2: Selection of representative biological terms and processes associated with the predicted target genes of motifs which activity was significantly **(a)** lower or **(b)** higher in promoted tumors relative to surrounding treated normal tissue, but that did not change in non-promoted tumors relative to surrounding non-treated normal tissue (**Supplementary Table S4**). Bars are colored according to motif to which the target genes are associated with. Bar height indicates significance of functional enrichment as it represents the $-\log10(P-\text{Value})$ of functional enrichment in the given biological term or process as obtained from the DAVID Bioinformatic Resource (Database for Annotation, Visualization and Integrated Discovery) [19, 20] version 6.7, sponsored by the National Institute of Allergy and Infectious Disease (NIAID), NIH.

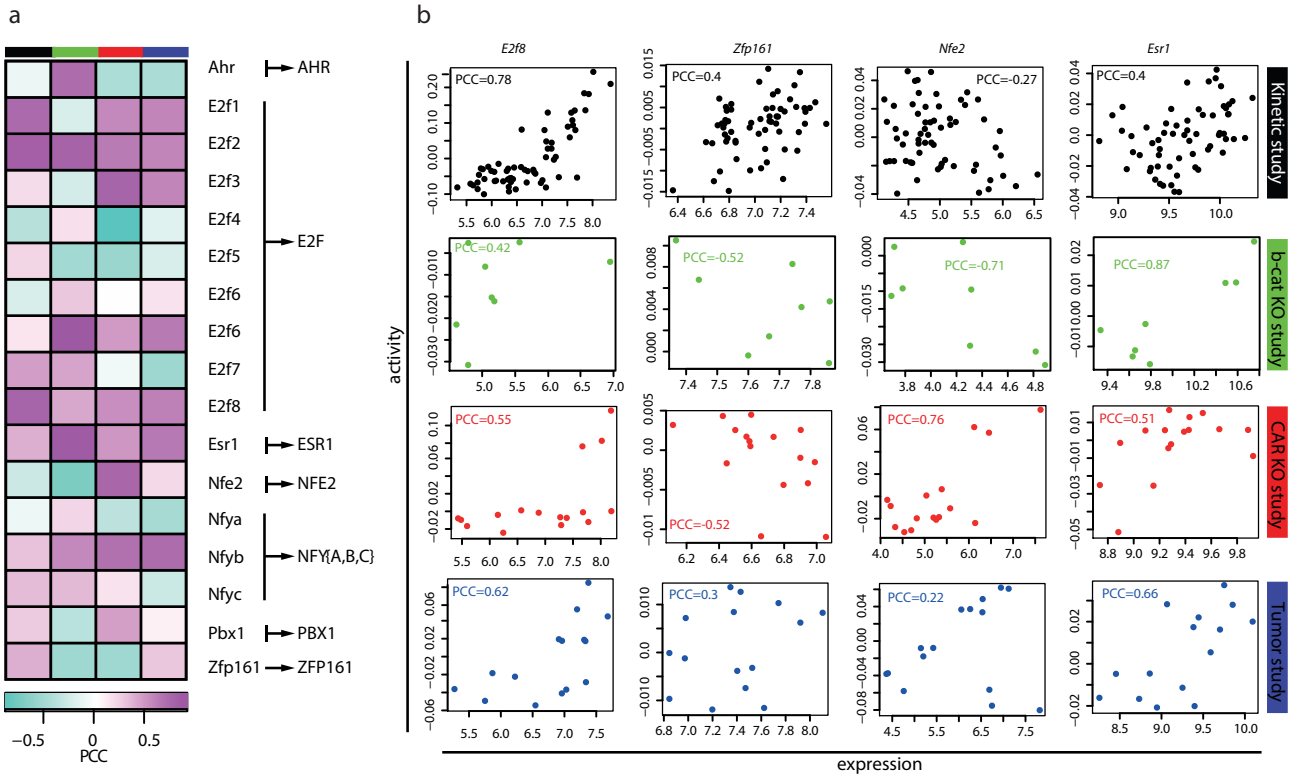

Figure S3: Correlation between motif activities and mRNA expression of cognate transcription factors. **(a)** Heatmap of the Pearson correlation coefficients (PCC) between the motif activities and mRNA expression profiles of associated TFs for a selection of TFs specifically dysregulated in promoted tumors. Each column corresponds to one of the 4 experimental data-sets (black = kinetic study, green = β-catenin KO study, red = CAR KO study and blue = tumor study) and PCC is indicated by color running from -1 (green), to 1 (purple). PCCs close to zero are colored white. **(b)** Scatter plots of motif activities against mRNA expression of associated TFs for a selection of 4 TFs. Each column of panels corresponds to one TF and each row of panels corresponds to one of the 4 experimental data-sets.

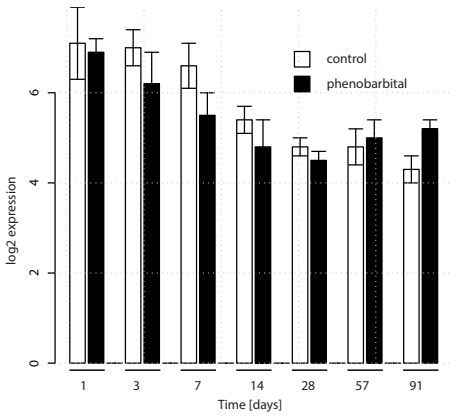

Figure S4: Alpha fetoprotein (Afp) gene expression in liver samples from 13 week kinetic data-sets as a surrogate gene of post-natal liver development termination. Gene expression is given as mean ±SD (n=3-5 animals per group). Open bars = control. Black bars = phenobarbital-treated samples.

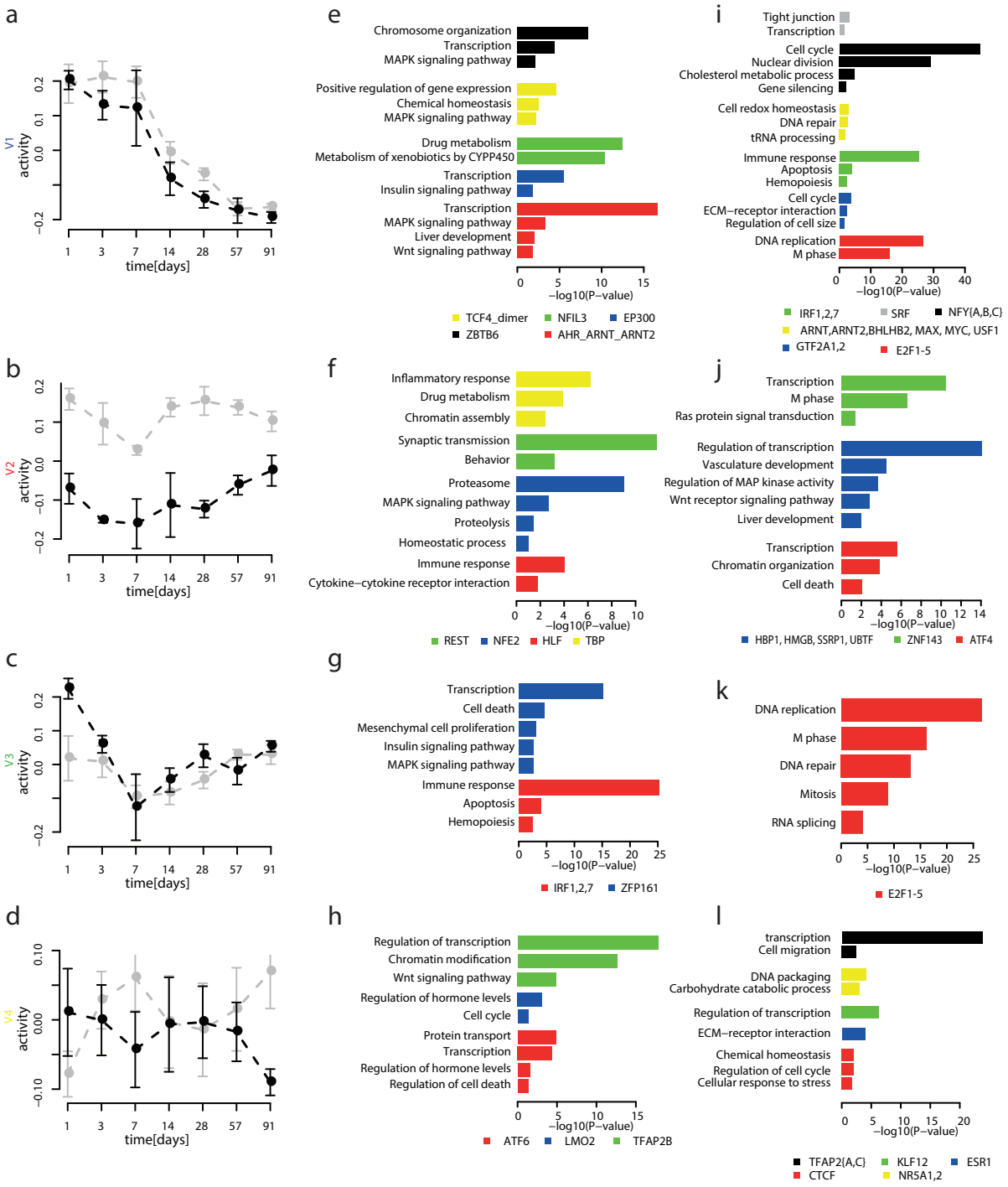

Figure S5: Gene Ontology and KEGG enrichment analysis of predicted targets for motifs underlying early PB-mediated transcriptional dynamics. **(a-d)** Plots of the activity profiles of the first four right singular vectors. **(e)-(l)** Selection of biological pathways and functional categories (Gene Ontology or KEGG) enriched among target genes of motifs that contribute/correlate negatively **(e-h)** or positively **(i-l)** to each of the singular vectors. Each color corresponds to one regulatory motif, indicated at the bottom of each panel, and the size of each bar corresponds to the significance ($-\log_{10}(p-\text{value})$) of the enrichment.

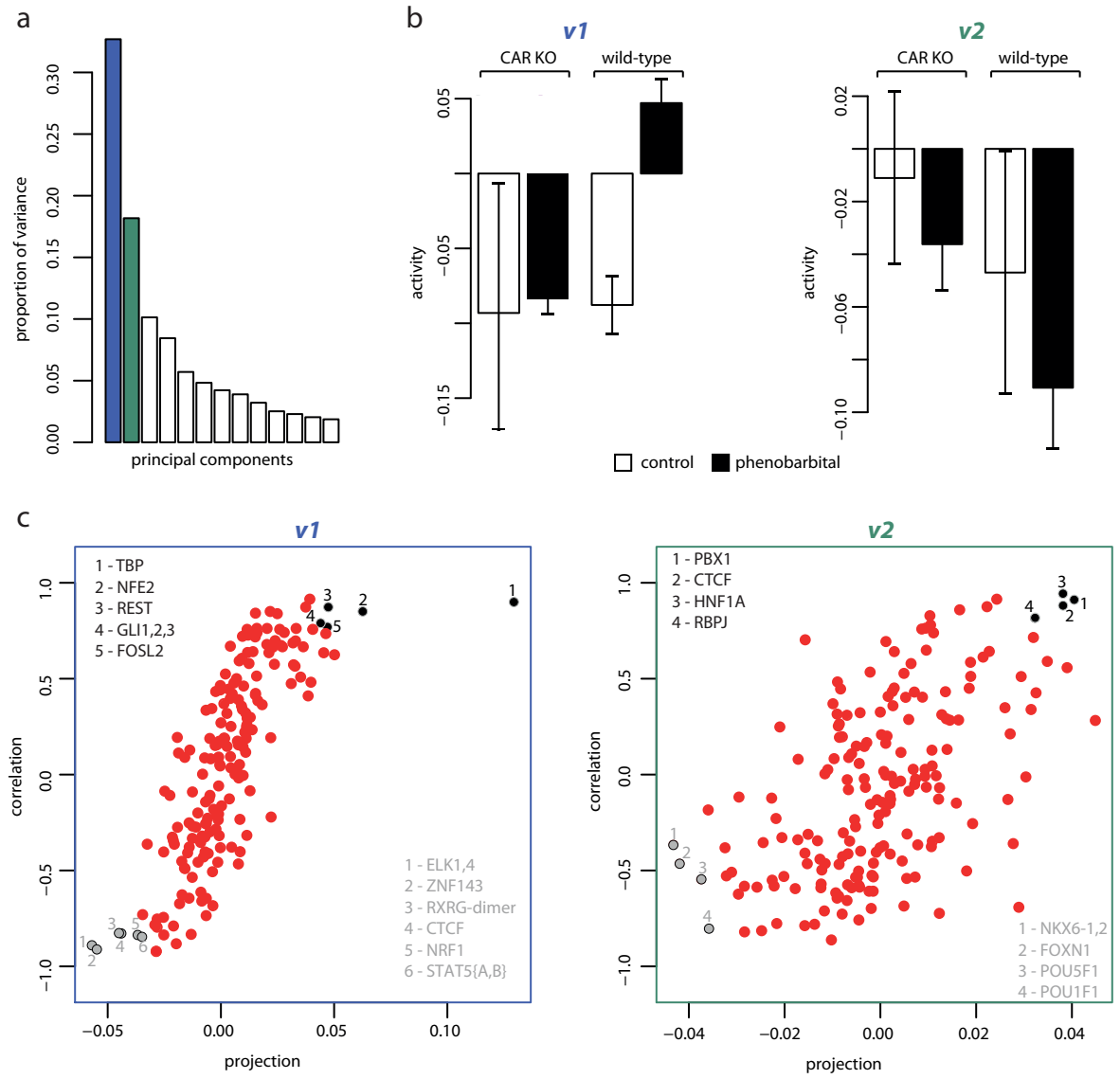

Figure S6: Singular Value Decomposition analysis of the activity matrix of the CAR KO data-set. **(a)** Proportion of the variance of the motif activity matrix. The first (blue bar) and second (green bar) components account for 33% and 18% respectively of the variance. **(b)** Barplot of the activity of the first two right singular vectors $v1$ and $v2$ in corresponding samples. White bars indicate activities for the control samples and black bars activities for the PB-treated samples. **(c)** Scatter plot of the correlations ρ_i and projections p_i of all motifs i with the first and second right singular vectors respectively. Grey and black dots depict negatively and positively selected motifs.

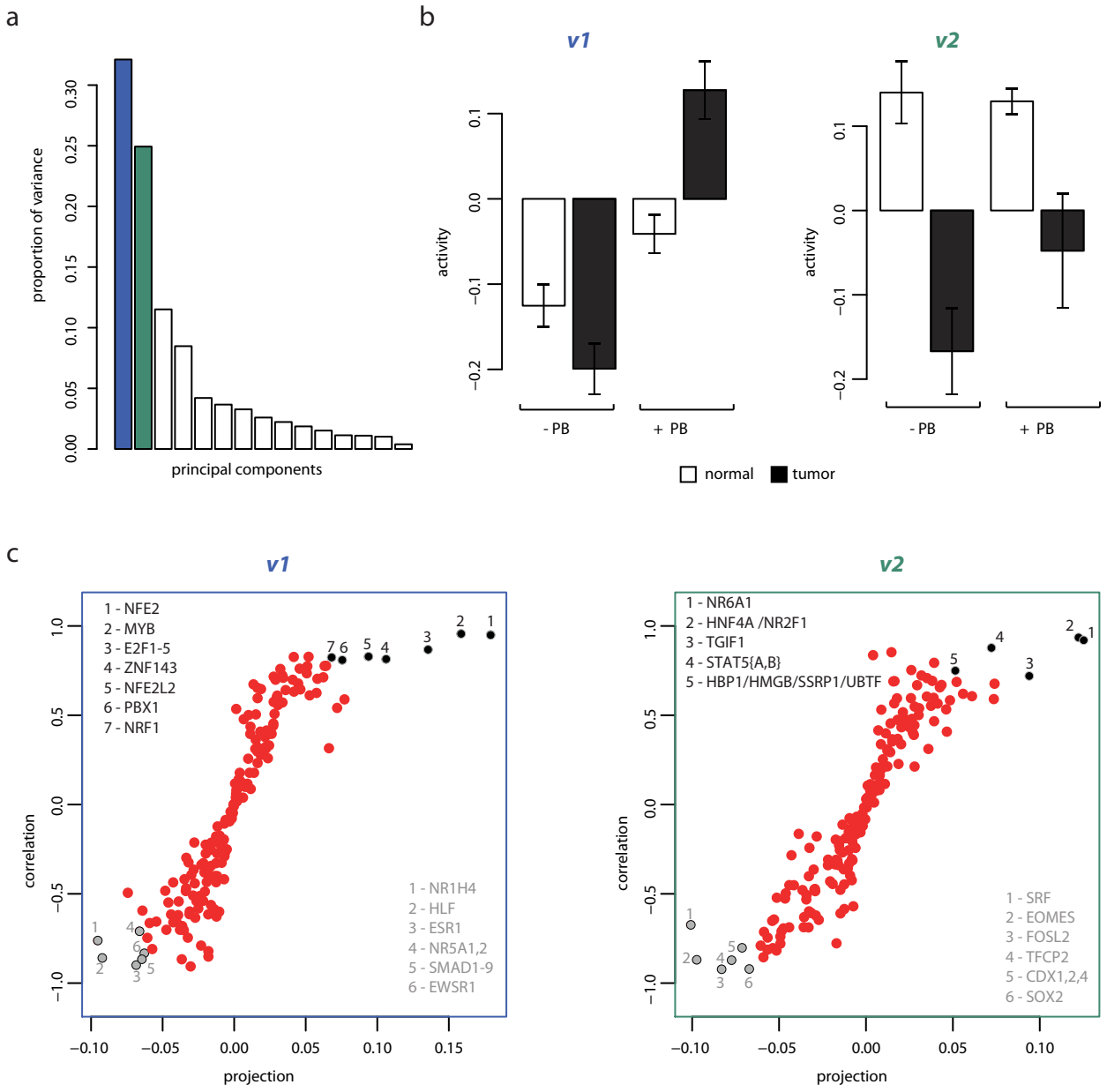

Figure S7: Singular Value Decomposition analysis of the activity matrix of the tumor data-set. **(a)** Proportion of the variance of the motif activity matrix captured by the first singular vectors. The first (blue bar) and second (green bar) components account for 32% and 25% respectively of the variance. **(b)** Barplot of the activity of the first two right singular vectors $v1$ and $v2$ across the corresponding samples. White bars indicate activities for the normal samples and black bars activities for the tumor samples. **(c)** Scatter plot of the correlations ρ_i and projections p_i of all motifs i with the first and second right singular vectors respectively. Grey and black dots depict negatively and positively selected motifs.

4 Abbreviations contained in Tables S1-S5

Tables S1-S5 contain motifs corresponding to specific groups that are

1. Table S1 - motifs associated with the first four singular vectors obtained from singular value decomposition (SVD) of the inferred motifs activity matrix from early kinetic study
2. Table S2 - motifs down-stream of CAR signaling
3. Table S3 - motifs dysregulated in both promoted and non-promoted tumors
4. Table S4 - motifs specifically dysregulated in promoted tumors
5. Table S5 - motifs down-stream of β-catenin signaling.

They are all formatted in the same way and their abbreviations are described in the following:

1. Representative motifs associated with the first four singular vectors obtained from SVD of the inferred motifs activity matrix from early kinetic study

   PC1 = first singular vector associated with liver maturation

   PC2 = second singular vector associated with constant xenobiotic response

   PC3 = third singular vector associated with transient mitogenic response

   PC4 = fourth singular vector associated with progressive xenobiotic response

   $+$ = motifs correlating positively with corresponding singular vector

   $-$ = motifs correlating negatively with corresponding singular vector

2. Z-value of motif significance that quantifies the significance of each motif in explaining the observed gene expression variation across the samples in the specified data-set

   S1 = kinetic data-set

   S2 = β-catenin KO data-set

   S3 = CAR KO data-set

   S4 = tumor data-set.

3. Z-values of differential motif activity that quantifies the evidence for a different regulatory activity of the motif between the tow following conditions

   $\mathbf{d}_i$ = PB-treated and control samples at corresponding time-point

   KO = knock-out and wild-type samples

   PB, wt = PB-treated and non-treated wild-type samples of the KO data-sets

   PB, ko = PB-treated and non-treated KO samples of the KO data-sets

   β-catenin = promoted tumors and treated surrounding normal tissue

   H-ras = non-promoted tumors and surrounding non-treated normal tissue.

5 Supplementary Tables

	Representative motifs				Motif Significance [z-value]				Differential Motif Activity													
									Kinetic study							b-catenin study	CAR study			Tumor study	
	PC1	PC2	PC3	PC4	S1	S2	S3	S4	d1	d3	d7	d14	d28	d57	d91	KO	KO	PB, *wt*	PB, *ko*	*β*-catenin	H-ras
E2F	+		+		3.7	1.8	2.3	2.7	**2.2**	**-0.7**	**-0.5**	**-1.2**	**-1.3**	**-0.8**	**0.3**	-2.5	-0.6	-0.7	-0.1	2.6	-0.3
IRF1,2,7	+		-		5.4	3.3	1.7	2.0	**-1.8**	**-0.8**	**0.2**	**-0.5**	**-1.7**	**0.6**	**-1.3**	3.7	0.6	1.8	0.3	0.6	0.1
SRF	+				3.0	1.3	0.7	1.9	**-2.4**	**-1.0**	**-1.0**	**-2.2**	**-0.6**	**0.4**	**-0.1**	1.7	0.2	-0.5	-1.0	1.0	1.9
NFY{A,B,C}	+				2.9	0.7	2.2	1.0	**1.3**	**-2.8**	**-3.0**	**-0.4**	**-0.3**	**-1.6**	**-1.3**	-1.0	0.2	-1.7	-0.2	1.7	0.8
GTF2A1,2	+				2.2	1.7	0.8	0.5	**3.1**	**1.2**	**-0.3**	**0.7**	**1.7**	**1.5**	**-0.5**	2.1	0.5	0.4	-0.5	-0.3	0.2
ARNT,ARNT2,BHLHB2,MAX,MYC,USF1	+				2.8	1.2	0.7	1.0	**2.2**	**-0.2**	**-2.4**	**-0.4**	**0.3**	**0.0**	**-0.8**	1.7	-0.9	-1.0	0.2	0.4	2.1
ZBTB6	-				2.2	0.4	1.1	1.3	**-1.8**	**-0.7**	**-0.1**	**-0.3**	**-0.2**	**-0.9**	**1.4**	-0.4	0.1	1.5	-0.5	-2.4	-1.5
TCF4-dimer	-				2.5	1.2	1.3	1.9	**0.1**	**0.4**	**0.5**	**0.0**	**-0.1**	**-2.1**	**-1.9**	-1.3	-0.1	-0.2	-0.2	-1.9	-3.2
NFIL3	-				2.4	1.5	1.3	0.9	**0.9**	**2.8**	**3.2**	**2.5**	**3.6**	**0.3**	**3.2**	-2.2	-0.9	0.9	0.7	0.6	-0.6
EP300	-				1.9	2.8	0.8	1.4	**0.1**	**0.4**	**0.9**	**0.1**	**0.1**	**0.4**	**-1.2**	-3.9	0.6	0.2	-0.1	-0.4	-2.3
AHR,ARNT,ARNT2	-				2.4	0.1	1.5	1.6	**-1.4**	**0.0**	**0.8**	**-0.6**	**-0.4**	**-2.0**	**-0.1**	0.2	0.1	-2.1	0.2	-1.9	-1.2
ZNF143		+			1.8	1.0	2.1	1.7	**0.0**	**-2.5**	**-2.3**	**-2.6**	**-2.1**	**-2.6**	**-2.1**	-1.3	-0.2	-3.0	-0.2	0.4	-0.1
HBP1,HMGB,SSRP1,UBTF		+			2.9	1.2	1.2	1.5	**-6.0**	**-4.1**	**-1.9**	**-2.4**	**-5.1**	**-2.6**	**-1.8**	1.7	0.9	0.7	0.8	-2.2	-0.9
ATF4		+			1.6	1.1	0.7	0.8	**-2.8**	**-2.4**	**-1.0**	**-1.5**	**-1.6**	**-3.3**	**-0.6**	1.5	0.1	-1.0	-0.5	0.6	0.8
TBP		-			2.7	0.3	2.2	0.5	**4.6**	**2.7**	**2.3**	**2.1**	**3.0**	**5.2**	**2.5**	0.3	-0.6	2.5	-0.4	-0.1	-0.5
REST		-			2.4	0.6	1.5	0.8	**2.3**	**2.8**	**4.4**	**2.0**	**4.0**	**1.8**	**0.7**	-0.5	0.0	1.8	0.1	0.5	-0.2
NFE2		-			1.7	1.6	3.0	2.6	**3.9**	**1.4**	**1.6**	**1.4**	**2.6**	**0.8**	**-0.4**	-2.2	-0.3	2.3	1.0	2.0	-1.4
HLF		-			1.8	0.5	1.3	1.4	**1.4**	**2.3**	**1.2**	**3.0**	**2.5**	**1.3**	**3.2**	0.7	0.4	2.6	-0.2	-1.4	1.1
ZFP161			-		1.6	0.3	1.0	1.6	**-2.0**	**-1.0**	**0.0**	**-1.1**	**-1.3**	**-0.7**	**1.9**	-0.2	-0.6	0.1	0.6	-1.6	0.7
TFAP2{A,C}				+	2.2	0.7	2.8	2.7	**-1.8**	**-1.9**	**-1.9**	**-1.1**	**-1.3**	**-1.4**	**-4.0**	1.0	0.8	-1.1	-0.5	-1.6	-3.7
NR6A1,2				+	1.2	2.0	2.1	2.3	**0.8**	**-0.1**	**0.8**	**0.4**	**0.4**	**0.5**	**-2.6**	1.9	1.7	-0.5	-1.8	-2.8	-2.0
KLF12				+	1.7	0.7	1.6	1.4	**0.4**	**-0.3**	**-1.4**	**0.6**	**0.8**	**0.8**	**-1.9**	1.0	0.2	-1.4	-0.2	-2.2	0.0
ESR1				+	1.7	1.9	1.5	1.9	**0.6**	**-0.7**	**-1.4**	**-0.1**	**-0.7**	**-0.8**	**-2.7**	2.6	1.4	0.0	-0.3	-2.4	0.5
CTCF				+	1.6	0.2	0.7	1.2	**-0.2**	**-2.3**	**-4.6**	**-1.5**	**-1.8**	**-1.1**	**-2.1**	0.3	0.7	-1.1	-0.5	0.6	2.8
TFAP2B				-	2.0	2.1	0.7	1.4	**-1.7**	**-1.1**	**0.5**	**-0.1**	**-0.5**	**-1.0**	**2.4**	-2.6	-0.7	-0.6	0.0	0.2	-0.6
LMO2				-	1.2	0.4	1.8	1.3	**-0.8**	**-0.3**	**-0.3**	**-0.3**	**0.3**	**2.6**	**2.0**	-0.4	-0.9	1.8	0.1	1.4	1.5
ATF6				-	1.1	0.6	0.7	1.4	**-1.8**	**-0.1**	**0.0**	**-0.2**	**0.6**	**0.8**	**2.5**	-0.9	0.0	0.7	-0.3	2.1	1.9

Table S1: Representative motifs of the first four singular vectors (explaining over 70% of the variance in the activity matrix) obtained from singular value decomposition of the inferred motifs activity matrix from early kinetic study, and underlying the early dysregulated biological pathways. Z-values of differential activity were computed as explained in Material and Method section of the main manuscript.

	Representative motifs				Motif Significance [z-value]				Differential Motif Activity													
									Kinetic study							b-catenin study	CAR study			Tumor study	
	PC1	PC2	PC3	PC4	S1	S2	S3	S4	d1	d3	d7	d14	d28	d57	d91	KO	KO	PB.*wt*	PB.*ko*	*β*-catenin	H-ras
Motifs down-stream of CAR signaling under physiological condition																					
NKX3-2					1.5	0.2	2.6	1.5	1.5	-0.3	-1.1	0.5	1.8	1.1	0.0	0.2	**2.7**	2.1	-1.3	1.7	2.6
FOX{F1,F2,J1}					1.6	0.2	1.9	1.1	0.8	2.0	1.2	1.3	0.5	2.4	0.8	0.2	**1.8**	2.2	-0.7	0.3	-0.4
NR5A1,2				+	1.2	2.0	2.1	2.3	0.8	-0.1	0.8	0.4	0.4	0.5	-2.6	1.9	**1.7**	-0.5	-1.8	-2.8	-2.0
ONECUT1,2					2.0	1.0	1.5	0.8	-1.5	-2.2	-2.4	-2.1	-1.3	-1.5	-1.6	-1.4	**-1.6**	-0.5	0.0	0.2	0.6
NKX2-2,8					1.5	1.2	2.2	1.4	-2.0	-0.9	0.9	-0.7	-0.8	0.4	1.7	1.7	**-1.6**	2.1	0.4	-0.7	-1.0
Motifs down-stream of CAR signaling under PB treatment																					
FOX{I1,J2}					1.6	0.2	2.2	0.4	0.5	1.7	1.6	1.2	1.3	2.9	0.7	0.3	1.1	**3.8**	**1.0**	-0.2	0.3
NFKB1,REL,RELA					1.8	0.7	2.2	2.1	0.5	1.9	-0.3	1.0	1.0	0.9	1.0	-0.5	-0.7	**2.7**	**0.3**	0.2	0.9
TBP				-	2.7	0.3	2.2	0.5	4.6	2.7	2.3	2.1	3.0	5.2	2.5	0.3	-0.6	**2.5**	**-0.4**	-0.1	-0.5
NFE2				-	1.7	1.6	3.0	2.6	3.9	1.4	1.6	1.4	2.6	0.8	-0.4	-2.2	-0.3	**2.3**	**1.0**	2.0	-1.4
PRDM1					1.4	0.5	2.2	1.1	0.0	1.1	0.7	0.9	2.1	0.9	0.8	0.7	0.8	**2.2**	**-0.2**	1.7	-0.2
FOX{F1,F2,J1}					1.6	0.2	1.9	1.1	0.8	2.0	1.2	1.3	0.5	2.4	0.8	0.2	1.8	**2.2**	**-0.7**	0.3	-0.4
NKX3-2					1.5	0.2	2.6	1.5	1.5	-0.3	-1.1	0.5	1.8	1.1	0.0	0.2	2.7	**2.1**	**-1.3**	1.7	2.6
NKX2-2,8					1.5	1.2	2.2	1.4	-2.0	-0.9	0.9	-0.7	-0.8	0.4	1.7	1.7	-1.6	**2.1**	**0.4**	-0.7	-1.0
REST			-		2.4	0.6	1.5	0.8	2.3	2.8	4.4	2.0	4.0	1.8	0.7	-0.5	0.0	**1.8**	**0.1**	0.5	-0.2
IRF1,2,7		+		-	5.4	3.3	1.7	2.0	-1.8	-0.8	0.2	-0.5	-1.7	0.6	-1.3	3.7	0.6	**1.8**	**0.3**	0.6	0.1
LMO2				-	1.2	0.4	1.8	1.3	-0.8	-0.3	-0.3	-0.3	0.3	2.6	2.0	-0.4	-0.9	**1.8**	**0.1**	1.4	1.5
FOSL2					0.7	0.7	1.6	1.4	-0.3	-0.4	0.0	0.4	-0.2	-0.1	1.1	0.9	0.1	**1.7**	**0.1**	1.2	2.4
RXR{A,B,G}					1.7	2.5	1.5	1.1	1.9	2.1	0.8	1.3	1.9	0.9	-0.6	3.5	0.0	**1.6**	**0.4**	-1.0	1.0
HOX{A4,D4}					2.2	0.7	1.7	1.1	0.1	0.5	0.8	1.3	-0.2	2.2	-0.4	0.8	1.2	**1.6**	**0.4**	-2.0	-0.7
GLI1-3					0.9	1.3	1.9	1.0	-1.5	-0.5	-0.9	-1.0	0.1	-0.5	-1.7	1.8	0.0	**1.5**	**-0.2**	0.4	-0.2
NFY{A,B,C}				+	2.9	0.7	2.2	1.0	1.3	-2.8	-3.0	-0.4	-0.3	-1.6	-1.3	-1.0	0.2	**-1.7**	**-0.2**	1.7	0.8
AHR,ARNT,ARNT2				-	2.4	0.1	1.5	1.6	-1.4	0.0	0.8	-0.6	-0.4	-2.0	-0.1	0.2	0.1	**-2.1**	**0.2**	-1.9	-1.2
CREB1					1.3	0.1	2.0	1.1	-0.2	-0.7	-1.7	-1.3	-1.2	-1.0	-0.4	-0.2	-0.6	**-2.1**	**0.6**	1.4	1.7
ELF1,2,4					2.2	0.5	1.9	2.6	-4.2	-2.7	-2.9	-2.5	-3.8	-1.8	-2.0	0.4	0.2	**-2.2**	**-0.3**	1.7	5.7
ELK1,4,GABP{A,B1}					2.3	0.6	1.9	1.1	0.1	-2.3	-1.4	-1.6	-0.8	-1.5	-1.2	-0.8	-0.1	**-2.6**	**-0.4**	0.8	1.3
ZNF143			+		1.8	1.0	2.1	1.7	0.0	-2.5	-2.3	-2.6	-2.1	-2.6	-2.1	-1.3	-0.2	**-3.0**	**-0.2**	0.4	-0.1
NRF1					1.5	0.5	1.8	1.8	1.1	-1.8	-1.3	-0.7	0.0	-0.6	-1.0	-0.6	-0.2	**-3.1**	**-0.3**	0.4	0.3
FOXD3					2.3	1.3	1.9	1.4	-3.0	-3.0	-1.5	-2.5	-4.1	-1.3	-0.9	0.9	0.0	**-3.6**	**-0.9**	-1.8	-0.7
Motifs differentially active upon PB treatment only in KO																					
SPI1					1.3	3.1	1.5	1.3	-3.9	-0.6	0.1	-0.5	-0.4	-2.0	1.2	4.3	-0.5	**-0.6**	**1.7**	-0.4	1.2
ZNF148					1.9	1.2	1.6	0.7	3.6	1.3	0.1	0.0	0.1	-0.2	-2.0	-1.7	0.8	**0.1**	**-2.0**	-0.1	-0.7
NR5A1,2				+	1.2	2.0	2.1	2.3	0.8	-0.1	0.8	0.4	0.4	0.5	-2.6	1.9	1.7	**-0.5**	**-1.8**	-2.8	-2.0
HNF4A,NR2F1,2					0.9	1.8	1.6	2.2	0.5	-0.4	-1.2	0.2	0.6	-0.2	-0.9	-2.2	0.2	**0.3**	**-1.6**	-1.5	-2.1

Table S2: Motifs which activities are significantly changing either 1) upon CAR KO in non-treated samples and thus potentially down-stream of CAR signaling under physiological condition, or under PB treatment 2) only in CAR wild-type samples and thus potentially down-stream of CAR signaling under PB treatment, or 3) only in CAR KO samples. Z-values of differential activity were computed as explained in Material and Method section of the main manuscript.

	Representative motifs				Motif Significance [z-value]				Differential Motif Activity												
									Kinetic study								b-catenin study	CAR study			Tumor study
	PC1	PC2	PC3	PC4	S1	S2	S3	S4	d1	d3	d7	d14	d28	d57	d91	KO	KO	PB,*wt*	PB,*ko*	*β*-catenin	H-ras
TFAP2{A,C}				+	2.2	0.7	2.8	2.7	-1.8	-1.9	-1.9	-1.1	-1.3	-1.4	-4.0	1.0	0.8	-1.1	-0.5	**-1.6**	**-3.7**
NR6A1					1.4	0.8	1.0	2.0	-0.4	0.4	0.9	-1.2	-1.0	-0.1	-2.0	-1.2	0.3	-1.0	-0.7	**-2.0**	**-3.5**
TCF4-dimer		-			2.5	1.2	1.3	1.9	0.1	0.4	0.5	0.0	-0.1	-2.1	-1.9	-1.3	-0.1	-0.2	-0.2	**-1.9**	**-3.2**
GATA6					0.9	0.8	1.6	1.7	-0.5	0.2	0.9	-0.4	-0.1	-1.4	-0.3	-0.9	0.5	0.1	-0.5	**-2.7**	**-2.3**
NR5A1,2				+	1.2	2.0	2.1	2.3	0.8	-0.1	0.8	0.4	0.4	0.5	-2.6	1.9	1.7	-0.5	-1.8	**-2.8**	**-2.0**
NR1H4					1.7	0.1	1.8	2.8	-1.4	-1.4	-0.4	-0.2	-0.2	-0.1	-1.6	0.0	0.9	0.1	-0.8	**-4.7**	**-1.7**
SOX{8,9,10}					1.0	0.6	1.0	1.6	-1.0	-0.7	0.1	-0.1	-1.2	0.7	-1.8	0.6	-0.8	-0.2	0.6	**2.0**	**2.5**

Table S3: Motifs which activities are significantly changing in promoted tumors relative to surrounding treated normal tissue, and in non-promoted tumors relative to surrounding non-treated normal tissue. These motifs are thus candidate regulators of liver tumorigenesis. Z-values of differential activity were computed as explained in Material and Method section of the main manuscript.

	Representative motifs				Motif Significance [z-value]				Differential Motif Activity												
									Kinetic study							b-catenin study	CAR study		Tumor study	
	PC1	PC2	PC3	PC4	S1	S2	S3	S4	d1	d3	d7	d14	d28	d57	d91	KO	KO	PB,*wt*	PB,*ko*	*β*-catenin	H-ras
NFE2		-			1.7	1.6	3.0	2.6	3.9	1.4	1.6	1.4	2.6	0.8	-0.4	-2.2	-0.3	2.3	1.0	**2.0**	**-1.4**
AHR,ARNT,ARNT2	-				2.4	0.1	1.5	1.6	-1.4	0.0	0.8	-0.6	-0.4	-2.0	-0.1	0.2	0.1	-2.1	0.2	**-1.9**	**-1.2**
E2F	+		+		3.7	1.8	2.3	2.7	2.2	-0.7	-0.5	-1.2	-1.3	-0.8	0.3	-2.5	-0.6	-0.7	-0.1	**2.6**	**-0.3**
ESR1				+	1.7	1.9	1.5	1.9	0.6	-0.7	-1.4	-0.1	-0.7	-0.8	-2.7	2.6	1.4	0.0	-0.3	**-2.4**	**0.5**
ZFP161			-		1.6	0.3	1.0	1.6	-2.0	-1.0	0.0	-1.1	-1.3	-0.7	1.9	-0.2	-0.6	0.1	0.6	**-1.6**	**0.7**
PBX1					1.3	0.4	1.8	1.8	1.2	-0.1	0.6	-0.7	0.0	0.8	-0.6	-0.5	0.9	-0.4	-0.8	**3.1**	**0.8**
NFY{A,B,C}		+			2.9	0.7	2.2	1.0	1.3	-2.8	-3.0	-0.4	-0.3	-1.6	-1.3	-1.0	0.2	-1.7	-0.2	**1.7**	**0.8**

Table S4: Motifs which activities are significantly changing in promoted tumors relative to surrounding treated normal tissue, but not in non-promoted tumors relative to surrounding non-treated normal tissue. These motifs are thus candidate regulators of tumor promotion. Z-values of differential activity were computed as explained in Material and Method section of the main manuscript.

	Representative motifs				Motif Significance [z-value]				Differential Motif Activity												
									Kinetic study							b-catenin study	CAR study			Tumor study	
	PC1	PC2	PC3	PC4	S1	S2	S3	S4	d1	d3	d7	d14	d28	d57	d91	KO	KO	PB,*wt*	PB,*ko*	*β*-catenin	H-ras
HNF1A					1.3	3.7	1.0	0.6	-2.7	-2.1	-0.9	-0.6	-1.8	-1.6	-0.9	**-4.7**	0.7	-1.0	-0.5	0.0	0.7
EP300	-				1.9	2.8	0.8	1.4	0.1	0.4	0.9	0.1	0.1	0.4	-1.2	**-3.9**	0.6	0.2	-0.1	-0.4	-2.3
POU6F1					0.8	2.2	0.7	0.6	-1.2	-0.8	-0.3	-1.7	-1.5	-0.6	0.2	**-2.9**	0.3	-0.3	0.8	-0.2	-0.5
LEF1,TCF7,TCF7L1,2					1.2	2.0	0.5	0.7	0.3	-0.3	1.2	-0.7	-1.4	0.2	0.0	**-2.8**	-0.1	-0.3	0.5	1.0	-0.6
TFAP2B				-	2.0	2.1	0.7	1.4	-1.7	-1.1	0.5	-0.1	-0.5	-1.0	2.4	**-2.6**	-0.7	-0.6	0.0	0.2	-0.6
ZNF384					2.0	1.8	0.5	1.0	-3.6	-2.0	-1.9	-2.1	-3.0	-1.9	0.8	**-2.6**	-0.9	-0.7	-0.1	1.3	-0.8
CDX1,2,4					3.0	1.9	1.2	2.5	2.1	2.8	0.1	2.3	2.3	-0.7	2.0	**-2.5**	-1.1	-0.6	0.1	4.0	0.9
E2F	+		+		3.7	1.8	2.3	2.7	2.2	-0.7	-0.5	-1.2	-1.3	-0.8	0.3	**-2.5**	-0.6	-0.7	-0.1	2.6	-0.3
NFE2L2					1.3	1.9	0.4	1.9	1.9	0.7	1.8	2.1	1.0	1.5	1.6	**-2.5**	0.0	0.7	-0.2	0.7	-2.2
RFX1-5,RFXANK,RFXAP					1.5	2.7	1.3	1.2	1.4	1.5	-1.0	1.1	0.7	0.6	-0.5	**-2.4**	-0.2	-0.3	-0.2	0.7	0.3
NFE2		-			1.7	1.6	3.0	2.6	3.9	1.4	1.6	1.4	2.6	0.8	-0.4	**-2.2**	-0.3	2.3	1.0	2.0	-1.4
ATF5,CREB3					1.2	2.0	0.4	0.6	-0.9	-1.2	0.1	-0.1	0.2	0.4	1.1	**-2.2**	0.2	-0.5	-0.6	-0.7	-0.3
HNF4A,NR2F1,2					0.9	1.8	1.6	2.2	0.5	-0.4	-1.2	0.2	0.6	-0.2	-0.9	**-2.2**	0.2	0.3	-1.6	-1.5	-2.1
NFIL3		-			2.4	1.5	1.3	0.9	0.9	2.8	3.2	2.5	3.6	0.3	3.2	**-2.2**	-0.9	0.9	0.7	0.6	-0.6
POU5F1,SOX2{dimer}					2.1	1.9	0.9	0.7	1.8	3.3	0.0	2.2	3.7	0.2	1.3	**-1.8**	-0.7	0.0	0.7	-0.3	-0.6
MYB					1.6	1.8	2.2	2.7	0.4	-0.3	0.3	-0.9	-0.5	-1.0	0.8	**-1.5**	-1.0	-1.1	-0.1	4.2	-0.4
GATA1-3					1.6	1.5	1.3	0.7	1.7	0.3	-2.5	1.1	0.9	0.8	-0.5	**1.7**	-0.3	1.5	1.6	-0.1	-0.2
FOXL1					1.9	1.6	1.6	1.0	-2.1	-1.7	-1.1	-3.1	-3.4	-1.6	-0.3	**1.8**	-0.1	-1.1	0.3	1.6	1.3
NR5A1,2				+	1.2	2.0	2.1	2.3	0.8	-0.1	0.8	0.4	0.4	0.5	-2.6	**1.9**	1.7	-0.5	-1.8	-2.8	-2.0
PAX2					1.8	1.6	1.0	0.7	-1.6	-3.0	-3.1	-3.0	-1.8	-2.0	-1.2	**2.0**	-0.1	0.2	-0.4	0.5	0.8
GTF2A1,2				+	2.2	1.7	0.8	0.5	3.1	1.2	-0.3	0.7	1.7	1.5	-0.5	**2.1**	0.5	0.4	-0.5	-0.3	0.2
TFAP4					1.5	2.0	1.2	0.5	-0.8	0.0	0.5	0.0	0.4	0.6	0.4	**2.3**	0.5	2.3	-0.3	-0.7	-0.1
MTF1					2.1	2.0	1.2	1.1	1.5	2.7	-0.3	0.5	1.1	2.0	-1.2	**2.6**	0.2	2.1	0.5	-0.7	-2.3
ESR1				+	1.7	1.9	1.5	1.9	0.6	-0.7	-1.4	-0.1	-0.7	-0.8	-2.7	**2.6**	1.4	0.0	-0.3	-2.4	0.5
KLF4					2.2	2.7	0.7	1.2	1.0	0.8	-1.4	-0.1	0.0	0.2	-1.4	**2.8**	0.5	-0.2	-0.1	1.1	-0.2
TFCP2					1.5	2.7	1.1	2.6	0.1	-1.7	-1.3	-0.2	-0.3	-0.1	0.2	**3.3**	0.8	0.2	0.5	0.7	3.7
RUNX1-3					1.1	2.6	1.2	0.9	1.3	-0.1	0.5	0.9	1.1	1.1	-0.4	**3.3**	1.0	1.5	-0.5	-0.8	-0.1
RXR{A,B,G}					1.7	2.5	1.5	1.1	1.9	2.1	0.8	1.3	1.9	0.9	-0.6	**3.5**	0.0	1.6	0.4	-1.0	1.0
IRF1,2,7			-		5.4	3.3	1.7	2.0	-1.8	-0.8	0.2	-0.5	-1.7	0.6	-1.3	**3.7**	0.6	1.8	0.3	0.6	0.1
TEAD1					1.0	2.9	1.0	0.9	-0.1	0.4	0.0	-0.3	-0.3	1.0	0.4	**4.1**	0.4	1.1	0.1	-0.3	2.0
FEV					2.5	3.0	0.6	0.8	1.4	0.7	-0.2	0.0	0.0	-0.1	-1.5	**4.3**	0.5	1.2	0.0	-0.1	-0.1
SPI1					1.3	3.1	1.5	1.3	-3.9	-0.6	0.1	-0.5	-0.4	-2.0	1.2	**4.3**	-0.5	-0.6	1.7	-0.4	1.2
FOS,FOS{B,L1},JUN{B,D}					1.5	3.4	0.4	1.5	-2.4	-1.9	-0.7	-1.8	-2.1	-1.3	-1.3	**4.8**	-0.3	0.1	0.7	0.2	2.3

Table S5: Motifs which activities are significantly changing upon β-catenin KO in non-treated samples and thus potentially down-stream of β-catenin signaling under physiological condition. Z-values of differential activity were computed as explained in Material and Method section of the main manuscript.

Affx	GS	Motifs	Kinetic study PCC	Kinetic study P-value	β-catenin study PCC	β-catenin study P-value	CAR KO study PCC	CAR KO study P-value	Tumor Study PCC	Tumor Study P-value
1450695_at	Ahr	AHR,ARNT,ARNT2	-0.09	5.1E-01	0.73	3.87E-02	-0.43	9.3E-02	-0.45	9.2E-02
1421721_a_at	Arnt	ARNT,ARNT2,BHLHB2,MAX,MYC,USF1	0.20	1.1E-01	-0.45	2.62E-01	0.43	9.9E-02	0.06	8.4E-01
1434028_at	Arnt2	ARNT,ARNT2,BHLHB2,MAX,MYC,USF1	0.07	5.9E-01	-0.19	6.54E-01	-0.05	8.6E-01	0.63	1.2E-02
1418025_at	Bhlhe40	ARNT,ARNT2,BHLHB2,MAX,MYC,USF1	-0.41	8.1E-04	0.26	5.26E-01	0.33	2.1E-01	0.33	2.3E-01
1423501_at	Max	ARNT,ARNT2,BHLHB2,MAX,MYC,USF1	0.01	9.3E-01	-0.07	8.65E-01	0.04	8.8E-01	-0.06	8.3E-01
1424942_a_at	Myc	ARNT,ARNT2,BHLHB2,MAX,MYC,USF1	0.11	4.1E-01	-0.20	6.36E-01	-0.02	9.5E-01	-0.28	3.2E-01
1448805_at	Usf1	ARNT,ARNT2,BHLHB2,MAX,MYC,USF1	0.10	4.2E-01	0.55	1.54E-01	0.25	3.4E-01	0.34	2.1E-01
1438992_x_at	Atf4	ATF4	0.21	1.0E-01	-0.23	5.76E-01	0.42	1.1E-01	-0.54	3.8E-02
1425927_a_at	Atf5	ATF5,CREB3	-0.40	1.1E-03	0.70	5.15E-02	0.28	2.9E-01	-0.51	5.3E-02
1419979_s_at	Creb3	ATF5,CREB3	0.45	2.1E-04	0.91	1.85E-03	-0.07	7.9E-01	-0.36	1.9E-01
1456021_at	Atf6	ATF6	0.26	4.5E-02	-0.46	2.56E-01	0.63	9.2E-03	0.72	2.2E-03
1449582_at	Cdx1	CDX1,2,4	0.06	6.4E-01	-0.25	5.58E-01	-0.32	2.3E-01	0.22	4.3E-01
1422074_at	Cdx2	CDX1,2,4	-0.01	9.3E-01	0.52	1.88E-01	-0.53	3.6E-02	-0.01	9.6E-01
1421552_at	Cdx4	CDX1,2,4	0.10	4.3E-01	-0.23	5.89E-01	0.15	5.7E-01	0.27	3.4E-01
1452901_at	Creb1	CREB1	0.59	4.5E-07	0.75	3.21E-02	-0.11	6.9E-01	-0.09	7.4E-01
1449042_at	Ctcf	CTCF	0.08	5.2E-01	0.25	5.49E-01	0.09	7.3E-01	0.30	2.7E-01
1418330_at	Ctcf	CTCF	0.21	1.0E-01	-0.12	7.83E-01	0.42	1.1E-01	-0.12	6.6E-01
1417878_at	E2f1	E2F	0.75	2.0E-12	-0.21	6.26E-01	0.58	1.9E-02	0.59	2.1E-02
1455790_at	E2f2	E2F	0.83	0.0E+00	0.80	1.83E-02	0.64	7.7E-03	0.59	2.0E-02
1434564_at	E2f3	E2F	0.20	1.3E-01	-0.23	5.89E-01	0.79	2.6E-04	0.59	2.0E-02
1451480_at	E2f4	E2F	-0.39	1.6E-03	0.18	6.70E-01	-0.88	8.7E-06	-0.16	5.7E-01
1447625_at	E2f5	E2F	0.23	6.7E-02	-0.50	2.09E-01	-0.56	2.6E-02	-0.21	4.5E-01
1448835_at	E2f6	E2F	0.15	2.4E-01	0.93	7.61E-04	0.49	5.2E-02	0.65	9.1E-03
1437187_at	E2f7	E2F	0.48	8.9E-05	0.45	2.66E-01	-0.06	8.3E-01	-0.53	4.0E-02
1436186_at	E2f8	E2F	0.78	1.0E-13	0.43	2.93E-01	0.55	2.8E-02	0.62	1.5E-02
1439319_at	Elf1	ELF1,2,4	0.18	1.7E-01	-0.40	3.26E-01	0.34	2.0E-01	0.23	4.0E-01
1428045_a_at	Elf2	ELF1,2,4	0.58	8.6E-07	0.57	1.39E-01	0.30	2.6E-01	0.33	2.2E-01
1421337_at	Elf4	ELF1,2,4	-0.14	2.7E-01	-0.21	6.21E-01	-0.37	1.6E-01	-0.42	1.2E-01
1446390_at	Elk1	ELK1,4,GABP{A,B1}	-0.02	8.7E-01	0.43	2.91E-01	-0.56	2.5E-02	-0.18	5.2E-01
1422233_at	Elk4	ELK1,4,GABP{A,B1}	-0.19	1.4E-01	-0.40	3.24E-01	-0.46	7.1E-02	-0.20	4.8E-01
1450665_at	Gabpa	ELK1,4,GABP{A,B1}	0.58	1.0E-06	0.03	9.38E-01	0.15	5.7E-01	0.05	8.7E-01
1436232_a_at	Gabpb1	ELK1,4,GABP{A,B1}	-0.13	3.1E-01	-0.53	1.76E-01	0.50	5.1E-02	0.17	5.5E-01
1460591_at	Esr1	ESR1	0.40	1.4E-03	0.87	4.48E-03	0.51	4.5E-02	0.66	7.9E-03
1425886_at	Fev	FEV	0.13	3.3E-01	-0.24	5.60E-01	0.35	1.9E-01	-0.16	5.7E-01
1423100_at	Fos	FOS,FOS{B,L1},JUN{B,D}	-0.14	2.9E-01	0.41	3.08E-01	0.55	2.9E-02	0.92	1.6E-06
1422134_at	Fosb	FOS,FOS{B,L1},JUN{B,D}	-0.08	5.5E-01	0.05	9.14E-01	0.29	2.7E-01	-0.03	9.2E-01
1417487_at	Fosl1	FOS,FOS{B,L1},JUN{B,D}	0.08	5.5E-01	-0.39	3.34E-01	0.74	9.3E-04	0.38	1.7E-01
1422931_at	Fosl2	FOSL2	-0.16	2.1E-01	-0.45	2.65E-01	0.63	8.8E-03	-0.34	2.1E-01
1434939_at	Foxf1	FOX{F1,F2,J1}	-0.57	1.2E-06	0.19	6.52E-01	-0.60	1.4E-02	-0.10	7.3E-01
1447562_at	Foxf2	FOX{F1,F2,J1}	0.38	2.2E-03	0.44	2.71E-01	0.15	5.8E-01	-0.05	8.7E-01
1425291_at	Foxj1	FOX{F1,F2,J1}	0.20	1.1E-01	-0.21	6.26E-01	-0.33	2.1E-01	-0.46	8.6E-02
1449458_at	Foxi1	FOX{I1,J2}	-0.32	1.0E-02	-0.05	9.16E-01	-0.15	5.8E-01	0.32	2.5E-01
1420374_at	Foxj2	FOX{I1,J2}	-0.11	4.0E-01	0.32	4.33E-01	0.23	3.9E-01	0.06	8.2E-01
1422210_at	Foxd3	FOXD3	-0.41	9.5E-04	-0.84	8.88E-03	0.26	3.2E-01	-0.35	2.1E-01
1423027_at	Foxl1	FOXL1	-0.03	8.1E-01	-0.68	6.33E-02	-0.47	6.7E-02	0.64	1.1E-02
1449232_at	Gata1	GATA1-3	-0.09	5.1E-01	-0.15	7.17E-01	0.20	4.5E-01	0.04	8.9E-01
1428816_a_at	Gata2	GATA1-3	0.02	8.9E-01	-0.15	7.15E-01	0.43	9.3E-02	0.36	1.8E-01
1448886_at	Gata3	GATA1-3	0.18	1.5E-01	-0.82	1.34E-02	0.05	8.6E-01	-0.71	2.8E-03
1425464_at	Gata6	GATA6	0.06	6.7E-01	-0.49	2.20E-01	-0.50	4.8E-02	-0.24	3.9E-01
1449058_at	Gli1	GLI1-3	0.11	3.9E-01	-0.19	6.56E-01	0.29	2.7E-01	0.62	1.3E-02
1446086_s_at	Gli2	GLI1-3	0.23	7.0E-02	-0.40	3.20E-01	-0.25	3.5E-01	-0.06	8.3E-01
1455154_at	Gli3	GLI1-3	0.18	1.5E-01	0.65	7.81E-02	0.01	9.8E-01	-0.01	9.6E-01
1450525_at	Gli3	GLI1-3	0.27	3.5E-02	0.17	6.92E-01	-0.54	3.0E-02	0.44	9.7E-02
1454631_at	Gtf2a1	GTF2A1-2	-0.60	2.5E-07	0.82	1.29E-02	-0.44	9.0E-02	-0.43	1.1E-01
1460367_at	Hbp1	HBP1,HMGB,SSRP1,UBTF	0.46	1.7E-04	-0.75	3.24E-02	0.26	3.4E-01	0.61	1.5E-02
1438307_at	Hmgb2	HBP1,HMGB,SSRP1,UBTF	0.13	3.2E-01	-0.65	7.81E-02	-0.41	1.1E-01	-0.51	5.3E-02
1416155_at	Hmgb3	HBP1,HMGB,SSRP1,UBTF	0.20	1.3E-01	-0.76	2.71E-02	-0.20	4.5E-01	-0.80	3.4E-04
1426788_a_at	Ssrp1	HBP1,HMGB,SSRP1,UBTF	0.09	4.9E-01	0.71	5.05E-02	-0.70	2.6E-03	-0.49	6.5E-02
1460304_a_at	Ubtf	HBP1.HMGB.SSRP1.UBTF	0.69	6.7E-10	-0.18	6.69E-01	-0.12	6.5E-01	0.17	5.5E-01
1434736_at	Hlf	HLF	-0.41	1.0E-03	0.10	8.20E-01	-0.68	4.0E-03	0.35	2.0E-01
1421234_at	Hnf1a	HNF1A	0.07	6.1E-01	0.45	2.63E-01	0.22	4.2E-01	-0.25	3.7E-01
1427000_at	Hnf4a	HNF4A,NR2F1,2	-0.02	9.0E-01	-0.43	2.86E-01	-0.35	1.8E-01	-0.64	1.1E-02
1418157_at	Nr2f1	HNF4A,NR2F1,2	-0.33	9.4E-03	-0.47	2.45E-01	-0.22	4.1E-01	0.40	1.4E-01
1416159_at	Nr2f2	HNF4A,NR2F1,2	0.45	2.3E-04	0.10	8.10E-01	-0.41	1.2E-01	0.69	4.1E-03
1427354_at	Hoxa4	HOX{A4,D4}	0.05	7.2E-01	0.16	7.12E-01	0.72	1.6E-03	0.20	4.8E-01
1450209_at	Hoxd4	HOX{A4,D4}	0.04	7.7E-01	0.40	3.29E-01	0.03	9.1E-01	0.11	7.0E-01
1448436_a_at	Irf1	IRF1,2,7	0.49	5.8E-05	0.10	8.17E-01	0.44	8.9E-02	0.87	2.4E-05
1418265_s_at	Irf2	IRF1,2,7	-0.20	1.2E-01	0.36	3.87E-01	-0.35	1.8E-01	-0.15	5.9E-01
1417244_a_at	Irf7	IRF1,2,7	0.71	1.0E-10	0.75	3.30E-02	0.61	1.2E-02	0.19	5.1E-01
1439846_at	Klf12	KLF12	-0.42	7.9E-04	0.57	1.41E-01	0.55	2.7E-02	0.04	9.0E-01
1417395_at	Klf4	KLF4	0.00	9.7E-01	0.75	3.36E-02	0.46	7.3E-02	-0.20	4.8E-01

Table S6: Pearson correlation coefficient (PCC) and associate P-values between motif activities and mRNA expression of cognate transcription factors in each data-sets - **part 1**. Part 2 in Table S7. Affx = probe-set ID from Affymetrix platform Mouse 430_2. GS = gene symbol. PCC = Pearson correlation coefficient.

Affx	GS	Motifs	Kinetic study PCC	Kinetic study P-value	β-catenin study PCC	β-catenin study P-value	CAR KO study PCC	CAR KO study P-value	Tumor Study PCC	Tumor Study P-value
1454734_at	Lef1	LEF1,TCF7,TCF7L1,2	-0.03	8.4E-01	0.30	4.73E-01	-0.25	3.5E-01	0.49	6.2E-02
1433471_at	Tcf7	LEF1,TCF7,TCF7L1,2	0.22	8.2E-02	0.46	2.52E-01	-0.26	3.3E-01	0.32	2.4E-01
1450117_at	Tcf7l1	LEF1,TCF7,TCF7L1,2	0.22	8.0E-02	-0.76	2.86E-02	0.25	3.5E-01	-0.43	1.1E-01
1426639_a_at	Tcf7l2	LEF1,TCF7,TCF7L1,2	0.32	1.2E-02	0.38	3.53E-01	0.45	8.3E-02	0.20	4.8E-01
1454086_a_at	Lmo2	LMO2	-0.01	9.7E-01	0.22	5.95E-01	-0.60	1.5E-02	-0.25	3.6E-01
1429170_a_at	Mtf1	MTF1	-0.55	3.4E-06	0.54	1.70E-01	-0.34	2.0E-01	-0.09	7.4E-01
1421317_x_at	Myb	MYB	-0.20	1.2E-01	-0.33	4.19E-01	-0.07	8.1E-01	-0.24	4.0E-01
1452001_at	Nfe2	NFE2	-0.27	3.5E-02	-0.71	4.74E-02	0.76	6.0E-04	0.22	4.3E-01
1457117_at	Nfe2l2	NFE2L2	-0.35	4.8E-03	-0.35	3.97E-01	-0.21	4.3E-01	0.22	4.2E-01
1418932_at	Nfil3	NFIL3	-0.07	5.7E-01	-0.31	4.49E-01	-0.55	2.7E-02	-0.10	7.3E-01
1427705_a_at	Nfkb1	NFKB1,REL,RELA	0.24	5.8E-02	-0.83	1.02E-02	0.48	5.8E-02	0.28	3.1E-01
1420710_at	Rel	NFKB1,REL,RELA	-0.44	3.1E-04	0.07	8.76E-01	0.18	5.2E-01	0.02	9.3E-01
1419536_a_at	Rela	NFKB1,REL,RELA	0.14	2.9E-01	0.55	1.62E-01	0.19	4.9E-01	0.13	6.3E-01
1427808_at	Nfya	NFY{A,B,C}	-0.08	5.2E-01	0.24	5.73E-01	-0.35	1.9E-01	-0.47	7.8E-02
1419266_at	Nfyb	NFY{A,B,C}	0.32	1.2E-02	0.58	1.36E-01	0.70	2.6E-03	0.73	2.1E-03
1448963_at	Nfyc	NFY{A,B,C}	0.35	4.9E-03	0.36	3.87E-01	0.17	5.3E-01	-0.28	3.1E-01
1421112_at	Nkx2-2	NKX2-2,8	-0.31	1.3E-02	0.35	3.96E-01	-0.09	7.3E-01	0.09	7.4E-01
1422284_at	Nkx2-9	NKX2-2,8	-0.46	1.9E-04	-0.10	8.21E-01	-0.17	5.3E-01	-0.19	4.9E-01
1421464_at	Nkx3-2	NKX3-2	0.21	9.5E-02	0.18	6.78E-01	-0.11	6.9E-01	-0.01	9.8E-01
1419105_at	Nr1h4	NR1H4	-0.16	2.0E-01	-0.19	6.56E-01	0.38	1.5E-01	0.66	7.0E-03
1421730_at	Nr5a1	NR5A1,2	0.26	4.4E-02	-0.20	6.29E-01	0.79	3.1E-04	0.37	1.7E-01
1449707_at	Nr5a2	NR5A1,2	0.07	5.8E-01	-0.29	4.79E-01	0.21	4.4E-01	-0.41	1.3E-01
1421515_at	Nr6a1	NR6A1	0.24	5.9E-02	0.37	3.69E-01	0.17	5.3E-01	0.38	1.6E-01
1424787_a_at	Nrf1	NRF1	0.55	4.2E-06	-0.36	3.80E-01	0.33	2.0E-01	-0.08	7.9E-01
1460044_at	Onecut2	ONECUT1,2	-0.27	3.3E-02	0.77	2.50E-02	0.41	1.2E-01	-0.29	2.9E-01
1428647_at	Pbx1	PBX1	0.29	2.2E-02	-0.37	3.63E-01	0.47	6.8E-02	0.10	7.4E-01
1416967_at	Sox2	POU5F1,SOX2{dimer}	0.45	2.7E-04	0.08	8.46E-01	0.20	4.6E-01	-0.23	4.1E-01
1452844_at	Pou6f1	POU6F1	0.51	2.6E-05	-0.68	6.22E-02	0.21	4.3E-01	0.00	1.0E+00
1420425_at	Prdm1	PRDM1	0.00	9.9E-01	-0.09	8.36E-01	0.36	1.7E-01	0.48	7.1E-02
1428227_at	Rest	REST	-0.64	2.3E-08	-0.62	9.92E-02	-0.54	3.1E-02	-0.06	8.3E-01
1436059_at	Rfx1	RFX1-5,RFXANK,RFXAP	-0.10	4.4E-01	0.57	1.43E-01	-0.36	1.7E-01	-0.43	1.1E-01
1442578_at	Rfx2	RFX1-5,RFXANK,RFXAP	0.23	7.3E-02	0.62	1.03E-01	0.55	2.6E-02	0.09	7.5E-01
1425413_at	Rfx3	RFX1-5,RFXANK,RFXAP	0.27	3.5E-02	-0.77	2.68E-02	0.49	5.4E-02	0.19	5.1E-01
1436931_at	Rfx4	RFX1-5,RFXANK,RFXAP	0.00	9.8E-01	0.78	2.20E-02	-0.59	1.6E-02	-0.32	2.5E-01
1425670_at	Rfxank	RFX1-5,RFXANK,RFXAP	0.44	3.9E-04	-0.11	7.94E-01	0.41	1.2E-01	0.23	4.0E-01
1455303_at	Rfxap	RFX1-5,RFXANK,RFXAP	0.53	1.1E-05	-0.08	8.59E-01	-0.34	2.0E-01	-0.10	7.1E-01
1440878_at	Runx1	RUNX1-3	0.05	6.8E-01	0.64	8.89E-02	0.15	5.7E-01	0.43	1.1E-01
1425389_a_at	Runx2	RUNX1-3	0.28	2.8E-02	-0.63	9.54E-02	0.29	2.8E-01	0.35	2.0E-01
1421467_at	Runx3	RUNX1-3	-0.06	6.7E-01	0.59	1.22E-01	0.65	5.9E-03	0.31	2.7E-01
1454773_at	Rxra	RXR{A,B,G}	0.27	3.7E-02	0.88	3.55E-03	-0.17	5.4E-01	0.43	1.1E-01
1416990_at	Rxrb	RXR{A,B,G}	-0.26	4.3E-02	0.09	8.27E-01	-0.04	8.9E-01	0.14	6.3E-01
1418782_at	Rxrg	RXR{A,B,G}	0.19	1.4E-01	0.24	5.68E-01	0.59	1.5E-02	-0.36	1.8E-01
1451689_a_at	Sox10	SOX{8,9,10}	0.02	8.5E-01	-0.49	2.21E-01	-0.66	5.3E-03	-0.59	2.1E-02
1435438_at	Sox8	SOX{8,9,10}	-0.06	6.7E-01	-0.23	5.82E-01	0.17	5.2E-01	0.40	1.4E-01
1451538_at	Sox9	SOX{8,9,10}	-0.24	5.5E-02	0.74	3.66E-02	-0.64	7.7E-03	-0.19	5.0E-01
1418747_at	Sfpi1	SPI1	-0.06	6.4E-01	0.08	8.59E-01	0.51	4.5E-02	0.25	3.6E-01
1418256_at	Srf	SRF	0.26	4.4E-02	0.22	6.08E-01	-0.02	9.5E-01	0.80	3.6E-04
1426470_at	Tbp	TBP	-0.07	6.0E-01	0.07	8.74E-01	-0.46	7.1E-02	-0.61	1.5E-02
1429556_at	Tead1	TEAD1	-0.02	9.0E-01	0.72	4.59E-02	0.40	1.2E-01	-0.13	6.5E-01
1436392_s_at	Tfap2c	TFAP{A,C}	0.07	6.1E-01	-0.70	5.47E-02	0.26	3.3E-01	0.45	9.2E-02
1426048_s_at	Tfap2a	TFAP2{A,C}	0.44	3.3E-04	0.36	3.75E-01	0.27	3.1E-01	0.47	7.4E-02
1435670_at	Tfap2b	TFAP2B	-0.36	4.6E-03	-0.02	9.67E-01	-0.61	1.1E-02	-0.05	8.5E-01
1418167_at	Tfap4	TFAP4	0.15	2.6E-01	-0.75	3.24E-02	-0.13	6.2E-01	-0.18	5.3E-01
1418159_at	Tfcp2	TFCP2	-0.30	1.7E-02	-0.46	2.48E-01	-0.22	4.0E-01	-0.21	4.5E-01
1455273_at	Zbtb6	ZBTB6	0.69	8.0E-10	0.52	1.83E-01	-0.39	1.3E-01	-0.01	9.7E-01
1420865_at	Zbtb14	ZFP161	0.40	1.2E-03	-0.52	1.86E-01	-0.52	3.8E-02	0.30	2.8E-01
1422599_s_at	Zfp143	ZNF143	0.41	8.4E-04	0.65	7.98E-02	0.74	1.1E-03	-0.14	6.1E-01
1436217_at	Zfp148	ZNF148	-0.73	1.7E-11	-0.46	2.56E-01	-0.28	3.0E-01	-0.72	2.3E-03
1438047_at	Zfp384	ZNF384	0.08	5.2E-01	-0.43	2.85E-01	0.08	7.6E-01	-0.19	4.9E-01

Table S7: Pearson correlation coefficient (PCC) and associate P-values between motif activities and mRNA expression of cognate transcription factors in each data-sets - **part 2**. Affx = probe-set ID from Affymetrix platform Mouse 430_2. GS = gene symbol. PCC = Pearson correlation coefficient.

References

[1] Lempiainen, H., et al. (2013 Feb) Identification of dlk1-dio3 imprinted gene cluster noncoding rnas as novel candidate biomarkers for liver tumor promotion. *Toxicol Sci*, **131**, 375–386.

[2] Tan, X., Behari, J., Cieply, B., Michalopoulos, G. K., and Monga, S. P. S. (2006 Nov) Conditional deletion of beta-catenin reveals its role in liver growth and regeneration. *Gastroenterology*, **131**, 1561–1572.

[3] Phillips, J. M., Burgoon, L. D., and Goodman, J. I. (2009 Aug) The constitutive active/androstane receptor facilitates unique phenobarbital-induced expression changes of genes involved in key pathways in precancerous liver and liver tumors. *Toxicol Sci*, **110**, 319–333.

[4] R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna Austria.

[5] Kauffmann, A., Gentleman, R., and Huber, W. (2009 Feb 1) arrayqualitymetrics–a bioconductor package for quality assessment of microarray data. *Bioinformatics*, **25**, 415–416.

[6] Septer, S., Edwards, G., Gunewardena, S., Wolfe, A., Li, H., Daniel, J., and Apte, U. (2012) Yes-associated protein is involved in proliferation and differentiation during postnatal liver development. *Am J Physiol Gastrointest Liver Physiol*, **302**, G493–503.

[7] Hart, S. N., Cui, Y., Klaassen, C. D., and Zhong, X.-b. (2009) Three patterns of cytochrome p450 gene expression during liver maturation in mice. *Drug Metab Dispos*, **37**, 116–21.

[8] Helin, K. (1998) Regulation of cell proliferation by the {E2F} transcription factors. *Current Opinion in Genetics & Development*, **8**, 28 – 35.

[9] Rodriguez, J. L., et al. (2007) Transcription of the mat2a gene, coding for methionine adenosyltransferase, is up-regulated by e2f and sp1 at a chromatin level during proliferation of liver cells. *Int J Biochem Cell Biol*, **39**, 842–850.

[10] Wu, L., et al. (2001 Nov 22) The e2f1-3 transcription factors are essential for cellular proliferation. *Nature*, **414**, 457–462.

[11] Gille, H., Sharrocks, A. D., and Shaw, P. E. (1992 Jul 30) Phosphorylation of transcription factor p62tcf by map kinase stimulates ternary complex formation at c-fos promoter. *Nature*, **358**, 414–417.

[12] Baena, E., Gandarillas, A., Vallespinos, M., Zanet, J., Bachs, O., Redondo, C., Fabregat, I., Martinez-A, C., and de Alboran, I. M. (2005 May 17) c-myc regulates cell size and ploidy but is not essential for postnatal proliferation in liver. *Proc Natl Acad Sci U S A*, **102**, 7286–7291.

[13] Goldsworthy, T. L., Goldsworthy, S. M., Sprankle, C. S., and Butterworth, B. E. (1994 May) Expression of myc, fos and ha-ras associated with chemically induced cell proliferation in the rat liver. *Cell Prolif*, **27**, 269–278.

[14] Cui, Y. J., Yeager, R. L., Zhong, X.-B., and Klaassen, C. D. (2009) Ontogenic expression of hepatic ahr mrna is associated with histone h3k4 di-methylation during mouse liver development. *Toxicol Lett*, **189**, 184–90.

[15] Chesire, D. R., Dunn, T. A., Ewing, C. M., Luo, J., and Isaacs, W. B. (2004 Apr 1) Identification of aryl hydrocarbon receptor as a putative wnt/beta-catenin pathway target gene in prostate cancer cells. *Cancer Res*, **64**, 2523–2533.

[16] Maher, J. M., Cheng, X., Slitt, A. L., Dieter, M. Z., and Klaassen, C. D. (2005) Induction of the multidrug resistance-associated protein family of transporters by chemical activators of receptor-mediated pathways in mouse liver. *Drug Metab Dispos*, **33**, 956–62.

[17] Alnouti, Y. and Klaassen, C. D. (2008) Tissue distribution, ontogeny, and regulation of aldehyde dehydrogenase (aldh) enzymes mrna by prototypical microsomal enzyme inducers in mice. *Toxicol Sci*, **101**, 51–64.

[18] Lahvis, G. P. and Bradfield, C. A. (1998) Ahr null alleles: distinctive or different? *Biochem Pharmacol*, **56**, 781–787.

[19] Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009 Jan) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*, **37**, 1–13.

[20] Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using david bioinformatics resources. *Nat Protoc*, **4**, 44–57.

Chapter 4

Phenobarbital Induces Cell Cycle Transcriptional Responses in Mouse Liver Humanized for Constitutive Androstane and Pregnane X Receptors

This chapter contains the second manuscript of this thesis which describes a study where human relevance of rodent humanized model in drug toxicity assessment is discussed in terms of gene expression data. The manuscript has been published in *Toxicological Sciences* in April 2014.

TOXICOLOGICAL SCIENCES 2014
doi: 10.1093/toxsci/kfu038
Advance Access publication January 4, 2014

Phenobarbital Induces Cell Cycle Transcriptional Responses in Mouse Liver Humanized for Constitutive Androstane and Pregnane X Receptors

Raphaëlle Luisier,[*] Harri Lempiäinen,[*] Nina Scherbichler,[*] Albert Braeuning,[†] Miriam Geissler,[†] Valerie Dubost,[*] Arne Müller,[*] Nico Scheer,[‡] Salah-Dine Chibout,[*] Hisanori Hara,[§] Frank Picard,[§] Diethilde Theil,[*] Philippe Couttet,[*] Antonio Vitobello,[*] Olivier Grenet,[*] Bettina Grasl-Kraupp,[¶,1] Heidrun Ellinger-Ziegelbauer,[‖,1] John P. Thomson,[‖‖] Richard R. Meehan,[‖‖,1] Clifford R. Elcombe,[‖‖‖,1] Colin J. Henderson,[#,1] C. Roland Wolf,[#,1] Michael Schwarz,[†,1] Pierre Moulin,[*] Rémi Terranova,[*,2] and Jonathan G. Moggs[*,1,2]

*Preclinical Safety, Novartis Institutes for Biomedical Research, CH-4057 Basel, Switzerland; †Department of Toxicology, Institute of Experimental and Clinical Pharmacology and Toxicology, University of Tübingen, 72074 Tübingen, Germany; ‡TaconicArtemis, Neurather Ring 1, Köln 51063, Germany; §DMPK, Novartis Institutes for Biomedical Research, CH-4057 Basel, Switzerland; ¶Institute of Cancer Research, University of Vienna, 1090 Vienna, Austria; ‖Department of Investigational Toxicology, Bayer Pharma AG, 42096 Wuppertal, Germany; ‖‖MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, Scotland; ‖‖‖CXR Biosciences Ltd, 2 James Lindsay Place, Dundee DD1 5JJ, UK; and #Medical Research Institute, University of Dundee, Ninewells Hospital & Medical School, Dundee DD1 9SY, UK

[1]MARCAR consortium member.

[2]To whom correspondence should be addressed at Preclinical Safety, Novartis Institutes for Biomedical Research, CHBS WKL135, Klybeckstrasse 141, CH-4057 Basel, Switzerland. Fax +41 61 3241027. E-mail: remi.terranova@novartis.com, jonathan.moggs@novartis.com.

Received November 7, 2013; accepted February 8, 2014

The constitutive androstane receptor (CAR) and the pregnane X receptor (PXR) are closely related nuclear receptors involved in drug metabolism and play important roles in the mechanism of phenobarbital (PB)-induced rodent nongenotoxic hepatocarcinogenesis. Here, we have used a humanized CAR/PXR mouse model to examine potential species differences in receptor-dependent mechanisms underlying liver tissue molecular responses to PB. Early and late transcriptomic responses to sustained PB exposure were investigated in liver tissue from double knock-out CAR and PXR (CAR^{KO}-PXR^{KO}), double humanized CAR and PXR (CAR^{h}-PXR^{h}), and wild-type C57BL/6 mice. Wild-type and CAR^{h}-PXR^{h} mouse livers exhibited temporally and quantitatively similar transcriptional responses during 91 days of PB exposure including the sustained induction of the xenobiotic response gene *Cyp2b10*, the Wnt signaling inhibitor *Wisp1*, and noncoding RNA biomarkers from the *Dlk1-Dio3* locus. Transient induction of DNA replication (*Hells*, *Mcm6*, and *Esco2*) and mitotic genes (*Ccnb2, Cdc20*, and *Cdk1*) and the proliferation-related nuclear antigen *Mki67* were observed with peak expression occurring between 1 and 7 days PB exposure. All these transcriptional responses were absent in CAR^{KO}-PXR^{KO} mouse livers and largely reversible in wild-type and CAR^{h}-PXR^{h} mouse livers following 91 days of PB exposure and a subsequent 4-week recovery period. Furthermore, PB-mediated upregulation of the noncoding RNA *Meg3*, which has recently been associated with cellular pluripotency, exhibited a similar dose response and perivenous hepatocyte-specific localization in both wild-type and CAR^{h}-PXR^{h} mice. Thus, mouse livers coexpressing human CAR and PXR support both the xenobiotic metabolizing and the proliferative transcriptional responses following exposure to PB.

Key Words: **nongenotoxic carcinogenesis; phenobarbital; liver; proliferation; humanized mice; CAR; PXR; transcription; cancer risk assessment.**

Phenobarbital (PB), an anticonvulsant commonly used for treatment of epilepsy and other seizures, promotes both spontaneous and chemically induced liver tumors in rodents (Becker, 1982; Whysner *et al.*, 1996) and has been widely used as a model compound for studying molecular mechanisms underlying rodent nongenotoxic hepatocarcinogenesis (Elcombe *et al.*, 2014; Lempiainen *et al.*, 2011; Phillips *et al.*, 2009a,b; Ross *et al.*, 2010; Thomson *et al.*, 2013; Yamamoto *et al.*, 2004). Murine liver tumor promotion by PB is dependent on the constitutive androstane receptor (CAR) and β-catenin (Huang *et al.*, 2005; Rignall *et al.*, 2011; Yamamoto *et al.*, 2004), and prolonged PB treatment selects for *Ctnnb1*- (encoding β-catenin) mutated tumors (Aydinlik *et al.*, 2001). CAR is required for the early PB-induced gene expression and DNA methylation changes that accompany murine hepatocyte hypertrophy and proliferative responses (Phillips *et al.*, 2009a; Ross *et al.*, 2010). PB regulates the nuclear localization of CAR (Kawamoto *et al.*, 1999) through an indirect mechanism involving inhibition of epidermal growth factor receptor (EGFR) signaling (Mutoh *et al.*, 2013). In addition to CAR, PB also activates the pregnane X receptor (PXR) (Lehmann *et al.*, 1998), which has overlapping functions with CAR to regulate xenobiotic metabolism and detoxification in liver (Tolson and Wang, 2010), and whose coactivation may enhance CAR-mediated hepatocyte proliferation (Shizu *et al.*, 2013).

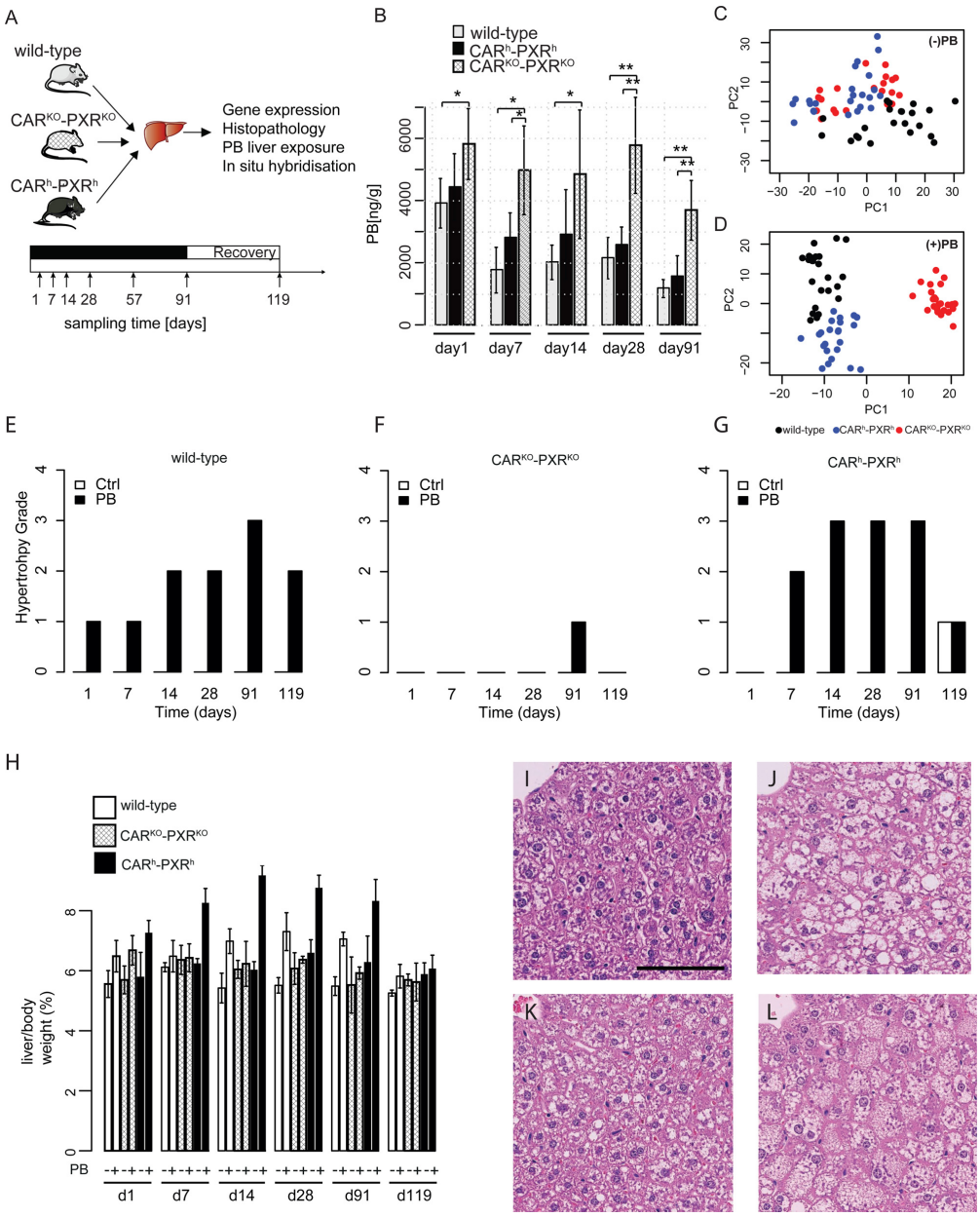

FIG. 1. (A) Experimental design of the kinetic phenobarbital study for molecular and phenotypic profiling. Male C57BL6 wild-type, CAR^{KO}-PXR^{KO}, and CAR^{h}-PXR^{h} mice were given *ad libitum* access to PB or vehicle (drinking water). Mice were sacrificed at indicated time points, livers were sampled and aliquots used for genome-wide gene expression (mRNA and miRNA), phenobarbital exposure, histopathology, and localization studies. (B) Effect of CAR/PXR KO and humanization on PB liver concentration in wild-type (gray bars), CAR/PXR humanized (black bars), and CAR/PXR KO (striped bars) male mice as obtained by LC-MS/MS. PB liver concentration (ng/g) is shown as a mean ± SD ($n = 5$ animals per group). Statistical significance is indicated by * (p-value < 0.05; Student's t-test), and ** (p-value < 0.01; Student's t-test). (C) PCA analysis performed on expression data from nontreated control samples, which cannot discriminate CAR^{KO}-PXR^{KO} from CAR^{h}-PXR^{h}, whereas PC2 discriminates wild-type from humanized and knock-out animals. (D) PCA analysis performed on expression data from treated samples allowed to discriminate the samples from the three strains: as PC1 discriminates CAR^{KO}-PXR^{KO} treated samples from wild-type and CAR^{h}-PXR^{h} treated samples, PC2 discriminates CAR^{h}-PXR^{h} treated samples from wild-type, and CAR^{KO}-PXR^{KO} treated samples. (E–G) Effect of phenobarbital on hypertrophy grade in wild-type, CAR^{KO}-PXR^{KO}, and CAR^{h}-PXR^{h} respectively over time. Severity grades were on a 1–4 scale and are expressed as median ($n = 5$). (H) Liver per body weight ratios (%) following PB treatment in wild-type, CAR^{KO}-PXR^{KO}, and CAR^{h}-PXR^{h} over time. Liver per body weight ratios are shown as a mean ± SD ($n = 5$ animals per group). (I–L) Hematoxylin and eosin (H&E) staining on liver sections from vehicle-treated WT (I) and PB-treated WT (J), PXR^{KO}-CAR^{KO} (K), and PXR^{h}-CAR^{h} (L) mice at day 91. Black bar: 100 μm.

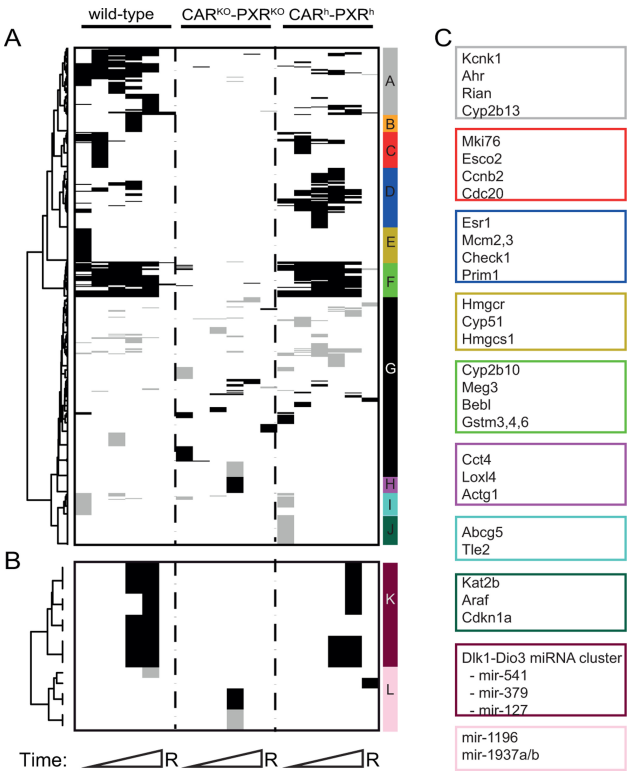

FIG. 2. Summary of PB-mediated differential mRNA (A) and miRNA (B) expression analysis in wild-type, CAR/PXR knock-out, and human CAR/PXR knock-in animals over time. Black horizontal bars = significantly upregulated transcripts; gray horizontal bars = significantly downregulated transcripts; white color indicates unchanged transcripts. A transcript is considered significantly up- or downregulated if $|\log_2 FC| > 0.53$ (corresponding to a FC > 1.5) and B.H. corrected *p*-value <0.01. Genes were hierarchically clustered by (1) computing Euclidean distance between genes from decision matrix and (2) applying Ward clustering algorithm. Only transcripts that are differentially expressed in at least one time point and in at least one mouse strain were included in the clustering analysis. Representative transcripts from each cluster are shown alongside the vertical color-coded cluster bars (A–L).

An increased understanding of mechanisms underlying chemical carcinogenesis has raised doubts regarding the appropriateness of extrapolating some rodent tumor findings to humans (Holsapple *et al.*, 2006). Based on epidemiological data in epileptics, there is no evidence of a specific role of PB in human liver cancer risk (IARC, 2001; Lamminpaa *et al.*, 2002; La Vecchia and Negri, 2013; Whysner *et al.*, 1996). Although prolonged treatment with PB does increase liver size in humans (Pirttiaho *et al.*, 1982), human hepatocytes are resistant to the ability of PB to increase cell proliferation (Hasmall and Roberts, 1999; Parzefall *et al.*, 1991) and inhibit apoptosis (Hasmall and Roberts, 1999).

The development of humanized mouse models provides a powerful approach for understanding pathways of human disease and to improve paradigms for the development of new drugs (Scheer *et al.*, 2013; Scheer and Wolf, 2013). They also have very significant potential in drug safety testing particularly in the light of the significant species differences in metabolism and toxicological response. Because the process of rodent nongenotoxic carcinogenesis is often mediated by nuclear receptors including CAR and PXR, humanized mouse models in which the endogenous mouse CAR/PXR genes have been replaced with human CAR/PXR genes have been used to explore the species specificity of PB-mediated hepatocellular responses. Humanized C57BL/6 CAR/PXR mice displayed induction of cytochrome P450's and hepatocellular hypertrophy but did not show hepatocyte proliferation following acute exposure to PB (Ross *et al.*, 2010). In contrast, PB-induced human CAR activation has been suggested to be associated with hepatocyte proliferation in an independent transgenic mouse model (Huang *et al.*, 2005). In this study, we have used a humanized (CAR^h-PXR^h) mouse model to examine potential species differences in receptor-dependent mechanisms underlying both early- and longer-term liver tissue transcriptomic responses to PB. We find that PB induces highly similar hepatic transcriptional programs in both wild-type and humanized CAR/PXR mice. This transcriptional response includes the upregulation of cell cycle genes, of the proliferative marker mKi67 and of *Dlk1-Dio3* locus noncoding RNAs that have recently been associated with cellular pluripotency (Lempiainen *et al.*, 2013; Stadtfeld *et al.*, 2010). These findings are discussed in the context of using humanized nuclear receptor mouse models to explore the human relevance of rodent nongenotoxic carcinogens.

MATERIALS AND METHODS

Ethics statement. The wild-type, null, and humanized CAR/PXR mouse 13-week time course study was performed in conformity with the Swiss Animal Welfare Law (specifically under the Animal License No. 2345 by Kantonales Veterinäramt Basel-Stadt (Cantonal Veterinary Office, Basel). The wild-type and humanized CAR/PXR mouse 4-week dose response study (Figs. 3B and 3C) was performed following University of Tübingen institutional guidelines.

Animal treatment and sample preparation. C57BL/6 male wild-type, knock-out CAR^{KO}-PXR^{KO} and humanized CAR^h-PXR^h mice (Scheer *et al.*, 2008,2010) were obtained from TaconicArtemis (Germany). For the 13-week time course study, 9–11 week-old mice (age selected to avoid the confounding effect of liver maturation observed in younger animals) were allowed to acclimatize for 5 days prior to being randomly divided into two treatment groups ($n = 5$ per time point). Phenobarbital (PB; free acid, >99.0%, Sigma, St Louis, MO, no. 04710, 0.05% (wt/vol) in drinking water) was administered to one group through *ad libitum* access to drinking water, as previously reported (Phillips *et al.*, 2009b; Lempiainen *et al.*, 2013; Thomson *et al.*, 2013). Mice were checked daily for ac-

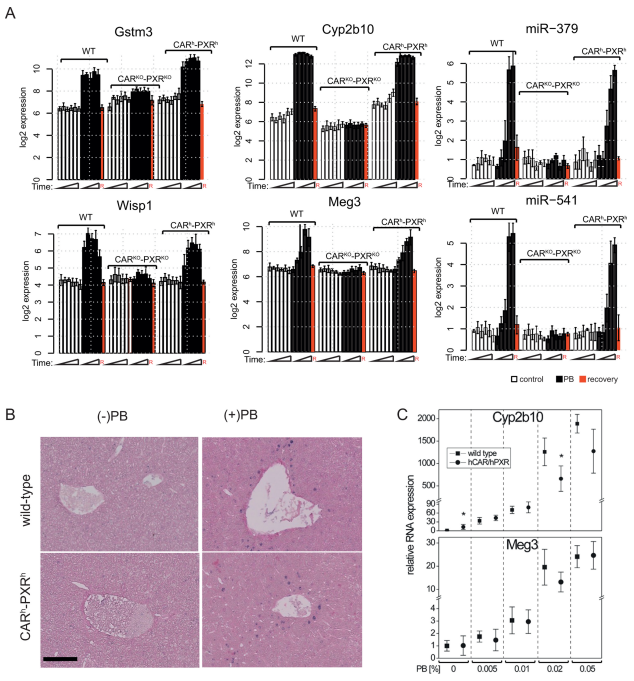

FIG. 3. PB mediates similar transcriptional changes in wild-type, CAR/PXR knock-out, and human CAR/PXR knock-in animals. (A) Expression of *Gstm3*, *Wisp1, Cyp2b10, Meg3*, miR-541, and miR-379 transcripts in control (open bars), treated (black bars) and recovery time point (red bars) male mice as determined using microarrays. Gene expression ($\log_2$) is given as a mean ± SD (n = 3–5 animals per group). (B) Expression and localization of *Meg3* in the liver. ISH of *Meg3* transcript (blue staining) in control (28 days) and PB-treated (28 days, and 0.05% (wt/vol)) livers from wild-type and humanized mice. Black bar = 100 μm. (C) Expression of *Cyp2b10* and *Meg3* obtained by RT-PCR in CAR^h-PXR^h mice and wild type controls, exposed to different concentrations of PB via the drinking water for 28 days. Mean ± SD (n = 4 mice per group) is shown relative to 18S rRNA expression. Statistically significant differences in PB-induced gene expression change between wild-type and CAR^h-PXR^h genotypes at a given dose are indicated by asterisks ($p < 0.05$; Student's t-test).

tivity and behavior and sacrificed on the indicated dates. For pharmacokinetics analysis, animals were anesthetized by inhalation of Isoflurane/O_2 and blood was taken from the orbital plexus in tubes containing EDTA. Blood samples were centrifuged at 3000 × g, plasma removed carefully from the EDTA containing tubes, snap frozen and stored at −80°C. For liver histopathology, one middle section of the median lobe (∼4–5 mm) was sampled, fixed in 10% neutral-buffered formalin for 48 h, processed, embedded in paraffin, and stained with hematoxylin/eosin. For molecular profiling, all remaining parts of the liver (incl. left and caudal parts) as well as the remaining median part of the liver were sampled. To ensure sample homogeneity for different molecular profiling methods, frozen liver samples were reduced to powder with CovarisCryoprep (Covaris Inc., Woburn, MA) system and aliquoted on dry ice. For the 4-week dose response study, mice (n = 4 per group) were treated with PB starting at 8 weeks of age. PB was administered via the drinking water (PB solution prepared freshly every fourth day) at concentrations of 0.005, 0.01, 0.02, and 0.05% (wt/vol). Mice were kept on a 12 h dark/light cycle and had accessed to food and water *ad libitum*. All animals were sacrificed between 9 and 11 A.M. to avoid circadian influences.

Phenobarbital exposure analysis in blood and liver tissue. Measurement of PB concentrations in plasma and liver were performed by Liquid Chromatography-Mass Spectrometry/Mass Spectometry (LC-MS/MS). At each necropsy time point, ∼0.2 ml plasma was sampled. Each pulverized liver sample, ca. 100 mg/tube, was diluted by adding 900 μl saline and mixed thoroughly to generate a homogenate. Twenty microliters of the homogenate was subjected to protein precipitation with 200 μl of internal standard solution (methanol) and 200 μl of 1:1 of methanol and acetonitrile. The supernatant after centrifugation was diluted by 20% of methanol and subjected onto LC-MS/MS determination. The separation on C18 column (Venusil ASB C18) was achieved by a gradient with 100% of H_2O and 100% of methanol with negative ion detection by turbo ion spray (API4000, Applied Biosystems). Data acquisition and peak integration were performed with software Analyst version 1.5.1 (Applied Biosystems). Student's t-test was used for statistical analysis of PB liver and plasma exposure. Differences were considered significant when p-value <0.05.

RNA isolation. Frozen liver samples were homogenized in TRIzol reagent (Invitrogen, Carlsbad, CA) and subsequently purified on a silica-gel-based-membrane (RNeasy, Qiagen, Venlo, Netherlands) according to the manufacturer's instructions. RNA quality was assessed by measuring the RIN (RNA Integrity Number) using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). RNA was stored at −80°C. miRNA was quantified using the Rediplate Ribogreen RNA quantitation kit (Life Technologies).

Affymetrix mRNA and microRNA GeneChip processing and gene expression data analysis. Affymetrix GeneChip Mouse Genome 430 2.0 arrays were used to profile mRNA expression in liver tissue of wild-type and transgenic mice (+/− PB) exposure. Affymetrix GeneChip miRNA 2.0 arrays with coverage of miRBase version 15 were used to profile miRNA expression in liver tissue of same samples. The array was able to detect the expression of 722 mouse mature miRNAs and 690 mouse pre-miRNA. Five biological replicates were used for each treatment group. Processing of GeneChip experiments was conducted as recommended by the manufacturer of the GeneChip system (Affymetrix, Santa Clara, CA).

All the analyses were performed with the R statistical package version 2.13 (2005) and Bioconductor libraries version 1.4.7 (R Core Team, 2013). Affymetrix CEL files were normalized using the Robust Multichip Average (RMA) implementation

of the algorithm available in R/Bioconductor (R Core Team, 2013). Quality metrics including scaling factors, average intensities, background intensities, RNA degradation, and raw *Q* values obtained from arrays prior and after normalization using Bioconductor's array QualityMetrics package (Kauffmann *et al.*, 2009) were within acceptable limits except for six chips which were removed from subsequent analyses. Affymetrix microRNA chips were preprocessed and normalized according to the Affymetrix miRNA QCTool manual. Briefly, the background control probes of the chips were grouped into bins of same dinucleotide Guanine-Cytosine (GC) content. The median signal of the background bin that matches with the GC content of the probe was then subtracted from the probe signal. The background corrected probes (for all probes on the chip including those of other species) were quantile normalized across chips, $\log_2$ transformed and summarized into probe sets with the median polish method as in standard RMA (Bolstad *et al.*, 2003). We floored all normalized signal values to 1.0.

Principal components analysis (PCA) was performed on the entire mRNA transcriptomic data set in order to characterize the overall structure of the data and identify major sources of gene expression variation. Independent PCA analyses were then performed on expression data from PB-treated and nontreated control samples in order to evaluate mouse genetic background effects (wild-type, CAR^{KO}-PXR^{KO}, and CAR^{h}-PXR^{h}) on PB treatment. Before statistical analysis, the transcriptomic data was filtered to remove 30% probes with lowest variation across samples and miRNA data was filtered to select for mouse probesets only. A two-way ANOVA model (treatment, time, and strain) was independently fitted to mRNA and miRNA data to assess statistical significance and linear contrasts at each time point using the Bioconductor's Limma package, which uses a Bayesian approach to better estimate the variance (Gentleman *et al.*, 2004). The Benjamini-Hochberg method was applied to correct for multiple comparisons (Benjamini and Hochberg, 1995). Probe sets with B.H. (Benjamini and Hochberg) corrected *p*-values <0.01 and absolute $\log_2$ fold changes above 0.53 (fold change >1.5 or <0.69) were considered differentially expressed. For cluster analysis, Euclidean distance was used as a similarity measure and the Ward method for agglomerative hierarchical clustering.

Real-time RT-PCR analyses. Total RNA was reverse transcribed by avian myeloblastosis virus reverse transcriptase (Promega). Relative RNA expression levels were analyzed on a LightCycler system (Roche) using the Fast Start DNA Master^{PLUS} SYBR Green Kit (Roche) according to the manufacturer's instructions, with the following primers: 18S rRNA_forward 5′-CGGCTACCACATCCAAGGAA-3′, 18S rRNA_reverse 5′-GCTGGAATTACCGCGGCT-3′, Cyp2b10_forward 5′-TACTCCTATTCCATGTCTCCAAA-3′, Cyp2b10_reverse 5′-TCCAGAAGTCTCTTTTCACATGT-3′, Meg3_forward 5′-GTCTTCCTGTGCCATTTGCT-3′, Meg3_reverse 5′-TTCATCAGTCAGTAGGTGGGTCT-3′. Gene expression values were calculated based on the crossing point differences and PCR efficiencies using the Pfaffl method (Pfaffl, 2001). 18S rRNA expression was used for normalization.

In situ *hybridization and immunohistochemistry. In situ* hybridization (ISH) and immunohistochemistry (IHC) analyses were conducted as described in the Supplementary Materials and Methods.

RESULTS

Similar PB Exposure and Hepatocellular Hypertrophic Responses in Wild-Type and Humanized CAR/PXR Mice

To explore potential species differences in CAR and PXR receptor-dependent mechanisms underlying liver tissue molecular responses to PB, a kinetic study (1, 7, 14, 28, and 91 days of treatment and a 4-week recovery group) was run in C57BL/6 wild-type, CAR^{KO}-PXR^{KO}, and CAR^{h}-PXR^{h} mice with *ad libitum* PB (0.05% wt/vol in drinking water) administration (Fig. 1A). Phenobarbital concentrations in plasma and liver as determined by liquid chromatography-mass spectrometry are shown in Figure 1B and Supplementary table 1. A significantly elevated PB plasma concentration was observed in all CAR^{KO}-PXR^{KO} mice compared with wild-type and CAR^{h}-PXR^{h} strains that may reflect a deficiency of their livers to metabolize PB (Fig. 1B). Similar PB liver concentrations that decreased after day 1 were observed in both wild-type and CAR^{h}-PXR^{h} mice indicating similar PB-induced metabolic activities in these two strains. The expression of mouse CAR and PXR transcripts was only detected in wild-type mouse livers (Supplementary fig. 4) consistent with the distinct nuclear receptor genotypes of CAR^{KO}-PXR^{KO} and CAR^{h}-PXR^{h} mice (Ross *et al.*, 2010).

The most striking PB-induced histopathological change observed in both wild-type and humanized CAR/PXR mice was hepatocellular hypertrophy starting from 7 days of PB treatment and increased in severity at later time points (Figs. 1E–L). Hypertrophy was primarily detected at perivenous hepatocytes in the central zone of the lobule (zone III) (Figs. 1I–L). PB did not induce hepatocellular hypertrophy in CAR^{KO}-PXR^{KO} mice (Figs. 1F and 1K) nor increase in liver to body weight ratio, consistent with previous reports (Huang *et al.*, 2005; Ross *et al.*, 2010). Very few mitotic figures were observed in any control or PB-treated animals. Hypertrophy evaluated by histopathology correlated well with increase in relative liver weight (Fig. 1H). Of note, PB induced higher grade of hepatocellular hypertrophy as well as higher relative liver weight in CAR^{h}-PXR^{h} animals compared with wild-type animals.

Distinct Transcriptional Patterns are Observed in Wild-Type, Null CAR/PXR, and Humanized CAR/PXR Mouse Livers in the Absence of PB Exposure

In the absence of PB exposure, wild-type and CAR^{h}-PXR^{h} mouse livers display some minor differences in their liver global transcriptome profiles including higher basal gene expression of *Cyp2b10* in livers from humanized CAR/PXR mice compared with wild-type mice (Supplementary table 4). PCA analysis of transcriptomic data from nontreated control samples discriminates wild-type from CAR^{h}-PXR^{h} and CAR^{KO}-PXR^{KO} mouse livers (Fig. 1C). This was confirmed by differential gene expression analysis of genes in control wild-type, CAR^{KO}-PXR^{KO}, and CAR^{h}-PXR^{h} mouse livers (Supplementary fig. 1 and Supplementary table 4) and is presumably due to unique constitutive hCAR and hPXR interactions with endogenous mouse gene regulatory proteins and gene regulatory DNA sequences. Additional factors that may contribute to these observed differences include mouse substrain genetic differences (the PXR/CAR double knock-out line is ~61% C57Bl/6J and 39% C57Bl/6N; the double humanized CAR/PXR line is ~78% C57Bl/6J and 22% C57Bl/6N; the wild-type comparator control mice were ~100% C57Bl/6J) (Scheer, TaconicArtemis, personal communication) and/or transgene-mediated perturbation of mouse hepatocyte chromatin architecture.

Mouse and Human CAR/PXR Mediate Similar and Reversible PB-Induced Mouse Liver Transcriptional Responses

PB-exposed wild-type and CAR^{h}-PXR^{h} mouse livers display distinct liver global transcriptome profiles compared with CAR^{KO}-PXR^{KO} mouse livers (Fig. 1D), consistent with PB predominantly inducing CAR- and PXR-mediated transcriptional responses (Phillips *et al.*, 2009a; Ueda *et al.*, 2002). Some degree of difference in PB-induced global liver transcriptional responses between wild-type and CAR^{h}-PXR^{h} mice was indicated by principal component 2 (PC2) discrimination within the PCA analysis (Fig. 1D). To further explore potential similarities and differences between mouse and human CAR/PXR-mediated regulation of hepatic genes upon PB-treatment, we compared mRNA and miRNA expression levels in livers from PB-treated wild-type, CAR^{KO}-PXR^{KO}, and CAR^{h}-PXR^{h} mice with their respective time-matched control samples. The differential expression of mRNAs and miRNAs in control wild-type, CAR^{KO}-PXR^{KO} and CAR^{h}-PXR^{h} mouse livers is summarized in Figures 2A and 2B and a detailed gene list for the corresponding clusters (A–L) is provided in Supplementary table 5. The most significantly upregulated genes in wild-type C57BL/6 mice upon PB-treatment included *Cyp2b10*, *Gstm3*, *Wisp1*, *Meig1*, *Abcc4*, *Cyp2b9*, *Cyp2c37*, *Prom1*, and *Gadd45b* (see Fig. 3A and Supplementary table 5) consistent with previous observations in PB-treated B6C3F1 mice (Lempiainen *et al.*, 2011, 2013; Phillips *et al.*, 2009a,b; Thomson *et al.*, 2013). Moreover, *Meg3* noncoding RNA and miRNAs associated with *the Dlk1-Dio3* imprinted cluster (miR-541 and miR-379), that have recently been associated with cellular pluripotency and proposed as novel candidate biomarkers for mouse liver nongenotoxic carcinogenesis (Lempiainen *et al.*, 2013), were also progressively upregulated upon PB-treatment (Figs. 2B and 3A). Importantly, similar quantitative and temporal responses for these PB-mediated molecular changes were also observed in humanized CAR/PXR mice whilst being absent in CAR/PXR null mice (Figs. 2B and 3A). Humanized CAR/PXR mice akin to wild-type mice also supported the PB-mediated induction of *Kcnk1*, a male-specific CAR-dependent transcriptional response to PB that has been associated with the attenuation of hepatic hyperplasia Saito *et al.,* 2013 . PB-induced *Meg3* induction in wild-type and humanized CAR/PXR mice was also confirmed to exhibit a very similar localization to perivenous hepatocytes in the central zone of the lobule (Fig. 3B).

The quantitative similarity between PB-mediated hepatic induction of *Cyp2b10* and *Meg3* in wild-type and CAR^{h}-PXR^{h} mice was further explored in a 4-week dose response study (Fig. 3C). Humanized CAR/PXR and wild-type mice of 8 weeks age were given four different concentrations of PB ranging from 0.005 to 0.05% (wt/vol) in drinking water. Both *Meg3* and *Cyp2b10* gene expression increased in a dose-dependent manner and appeared to reach a plateau at around 0.02% suggesting saturation (Fig. 3C).

Some differences in the magnitude and timing of liver transcriptional responses were observed between humanized CAR/PXR and wild-type mice. One group of PB-induced gene expression changes was observed to be more prominent in wild-type than CAR^{h}-PXR^{h} mice (see Fig. 4A for examples and Supplementary table 6 for a detailed list of genes which exhibit significant differences in PB-mediated gene induction between humanized and wild-type mice in at least one time point) and an additional distinct group of PB-induced gene expression changes was observed to be more prominent in CAR^{h}-PXR^{h} than the wild-type mice (e.g., see Fig. 4B). To assess whether PB exposure differences might contribute to the observed differential PB-induced expression of these genes, we performed linear modeling between gene expression and liver exposure but did not observe any significant effects (Supplementary fig. 2).

To determine the potential reversibility of molecular responses induced by 13 weeks of PB treatment, microarray experiments using liver samples from 13-week PB-treated mice followed by 4 weeks recovery (wild-type, CAR^{KO}-PXR^{KO}, and CAR^{h}-PXR^{h}) were compared with liver samples of nontreated mice from the same strains at the 119-day time point. Differential gene expression analyses revealed that only six genes (representing 1% of differentially expressed genes over time) maintained a qualitatively consistent residual differential expression in humanized and/or wild-type mice (*Serpinb1a*, *Nebl*, *Cyp2b13*, *Gna14*, *Gm20265*, and *AI132709*) after 4 weeks of recovery (see Supplementary table 7 for complete gene list) suggesting either residual CAR/PXR activity associated with positive gene regulation after removal of PB treatment, longer-term stability of this subset of PB-induced mRNAs in mouse liver or

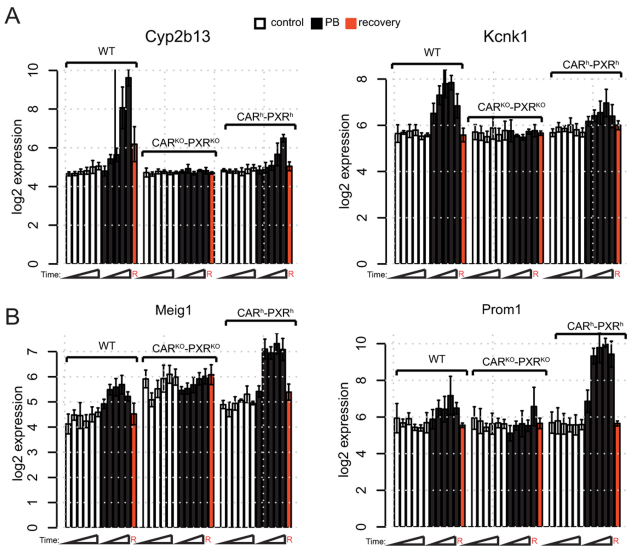

FIG. 4. Temporal and quantitative differences in PB-mediated transcriptional responses in wild-type versus humanized CAR/PXR mouse livers. (A) PB-mediated gene induction of *Cyp2b13* and *Kcnk1* is more prominent in wild-type than humanized animals. (B) PB-mediated gene induction of *Meig1* and *Prom1* is more prominent in humanized animals than wild-type. Expression ($\log_2$) in control (open bars), treated (black bars), and recovery time point (red bars) male animals is given as mean ± SD ($n = 3$–5 animals per group).

long-lasting epigenetic changes associated with their gene regulatory regions. Importantly, the 119-day expression levels for all of the above residual differentially expressed genes was significantly lower than expression levels observed following 91 days of PB treatment suggesting a slow return to normal basal gene expression levels. A small number of genes were uniquely differentially expressed at the recovery time point in humanized or knock-out mice (Supplementary table 7), including a number of inflammatory genes.

Human CAR and PXR Support Mouse Liver Transcriptional Upregulation of DNA Replication, Cell Cycle and Mitotic Genes upon PB Exposure

Further analysis of the different clusters associated with PB-mediated differential expression (Fig. 2A) revealed a transient CAR/PXR-dependent cell cycle response after 1 and 7 days of PB treatment characterized by upregulation of genes associated with DNA replication, cell cycle and mitosis (Supplementary tables 2 and 3; Figs. 5A and 5B). Importantly, this PB-mediated cell cycle transcriptional response was also observed in humanized CAR/PXR mice whilst being absent in CAR/PXR null mice. Further analysis of gene functions associated with this cluster by mapping to cell cycle phases suggests that PB supports all phases of cell cycle progression at the transcriptional level from S-phase entry to cytokinesis (Fig. 6).

DISCUSSION

This study demonstrates for the first time that mouse livers expressing only the human versions of CAR and PXR can support PB-induced xenobiotic and proliferative responses at the transcriptional level. Wild-type and CAR^h-PXR^h mouse livers exhibited temporally and quantitatively similar transcriptional responses during 91 days of PB exposure including the sustained induction of the xenobiotic response gene *Cyp2b10*, the Wnt signaling inhibitor *Wisp1* and noncoding RNA biomarkers from the *Dlk1-Dio3* locus (Lempiainen *et al.*, 2013). Importantly, mouse livers expressing human CAR and PXR also supported PB-mediated transient DNA replication (*Hells*, *Mcm6*, and *Esco2*), cell cycle (*Ccnb2*, *Cdc20*, and *Cdk1*) and proliferation-related nuclear antigen *Mki67* transcriptional responses consistent with hepatocyte proliferation. Our data are consistent with a previous report that PB (0.05%; 1 week) induced human CAR activation in a transgenic mouse model associated with hepatocyte DNA replication based on increased ploidy and PCNA protein expression (Huang *et al.*, 2005). Another study in humanized C57BL/6 CAR/PXR mice concluded that PB-induced cytochrome P450's and hepatocellular hypertrophy but not hepatocyte DNA replication (based on BrdU incorporation) following acute exposure to PB (4 days; ip 80 mg/kg) (Ross *et al.*, 2010). However, a more recent follow-up 7-day study using an alternate dietary route of administration to generate higher systemic exposure revealed that CAR^h-PXR^h mice can support PB-mediated DNA replication (based on BrdU incorporation) although at higher systemic exposures relative to wild-type mice (Elcombe, personal communication). The measurement of PB-induced mouse liver proliferation is confounded by changes in a heterogenous liver parenchymal cell population that include both mononuclear and binucleated cells containing multiples of the diploid component of DNA (known as polyploidy) (Bohm and Noltemeyer, 1981; Bursch *et al.*, 2004; Gonzales *et al.*, 1998) that may result from incomplete cytokinesis, endoreplication, or a combination of both (Gentric *et al.*, 2012a,b; Styles, 1993). Thus, increase in either Ki67 mRNA expression, Ki-67/PCNA IHC staining, or BrdU incorporation may reflect the increased DNA content of a polyploid hepatocyte rather than DNA replication associated with complete cell division cycles and hyperplasia. However, our observation of progressive PB-induction of genes that drive S-phase, mitosis, and cytokinesis in wild-type and humanized CAR/PXR mouse livers (Fig. 6) supports a proliferative response. Furthermore, we did not observe PB-mediated changes in the expression of genes implicated in driving polyploidization of hepatocytes such as *E2f7/8* and *c-Myc* (see Supplementary fig. 3) (Pandit *et al.*, 2012, 2013).

Species differences in CAR activation by direct ligands such as the human CAR-selective CITCO (Auerbach *et al.*, 2005) and mouse CAR-selective TCPOBOP (Nims *et al.*, 1993) are thought to be in part due to differences in ligand-binding domain protein structure. Murine and human CAR and PXR lig-

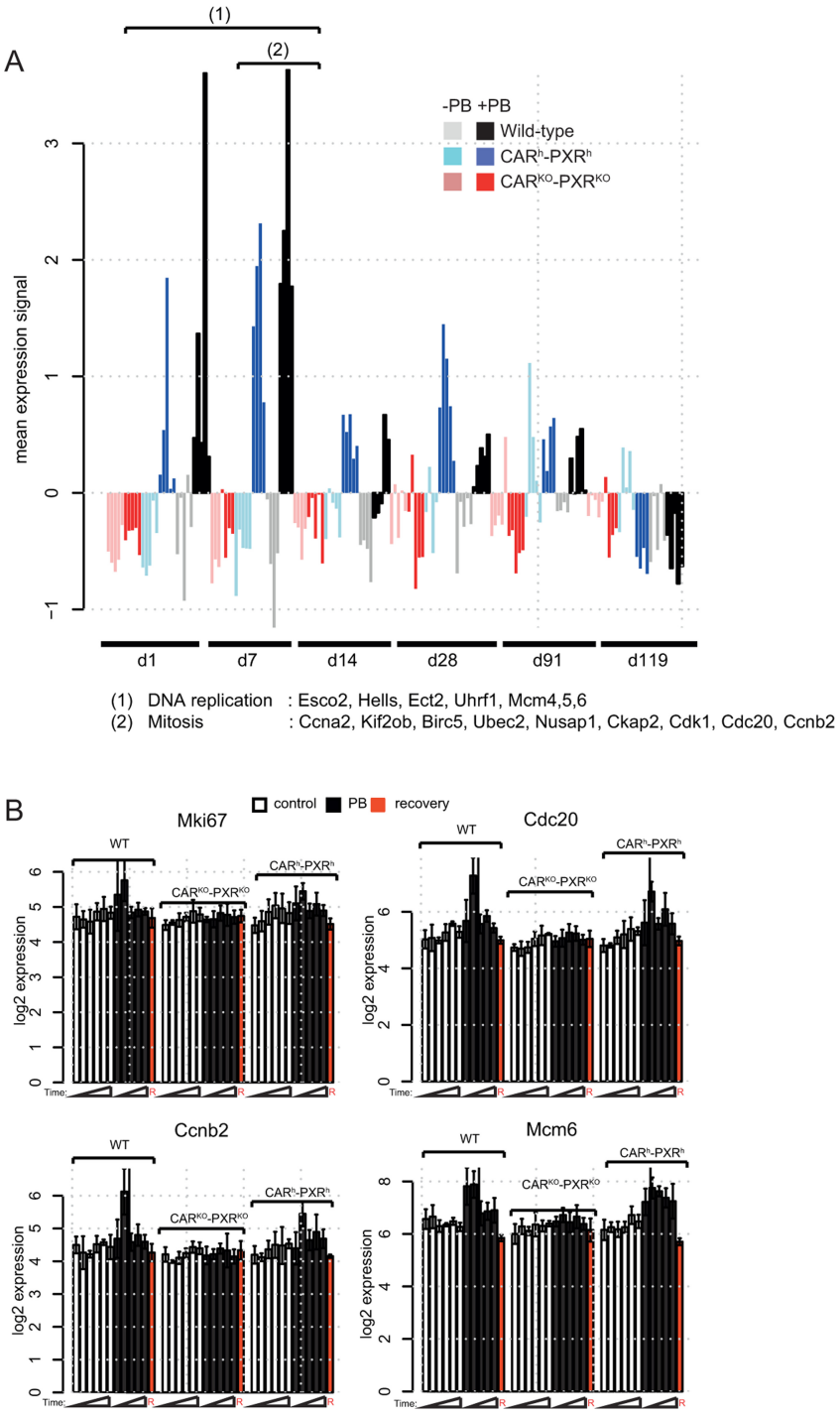

FIG. 5. PB mediates cell cycle transcriptional responses in wild-type and humanized CAR/PXR mouse livers. (A) Median expression per animal of a cluster of genes enriched for DNA replication and cell cycle functions that were differentially expressed upon PB treatment between day 1 and day 7 (cluster C from Fig. 2A). PB-mediated entry in S-phase after one day is supported by a subset of DNA replication genes significantly upregulated from day 1 (1) whereas mitosis and cytokinesis are supported by a subset of cell cycle regulatory genes significantly upregulated around day 7 (2). (B) Expression of *Mki67*, *Cdc20*, *Ccn20*, and *Mcm6* determined using microarrays. Expression in control (open bars), treated (black bars) and recovery time point male mice is given as mean ± SD ($n = 3$–5 animals per group).

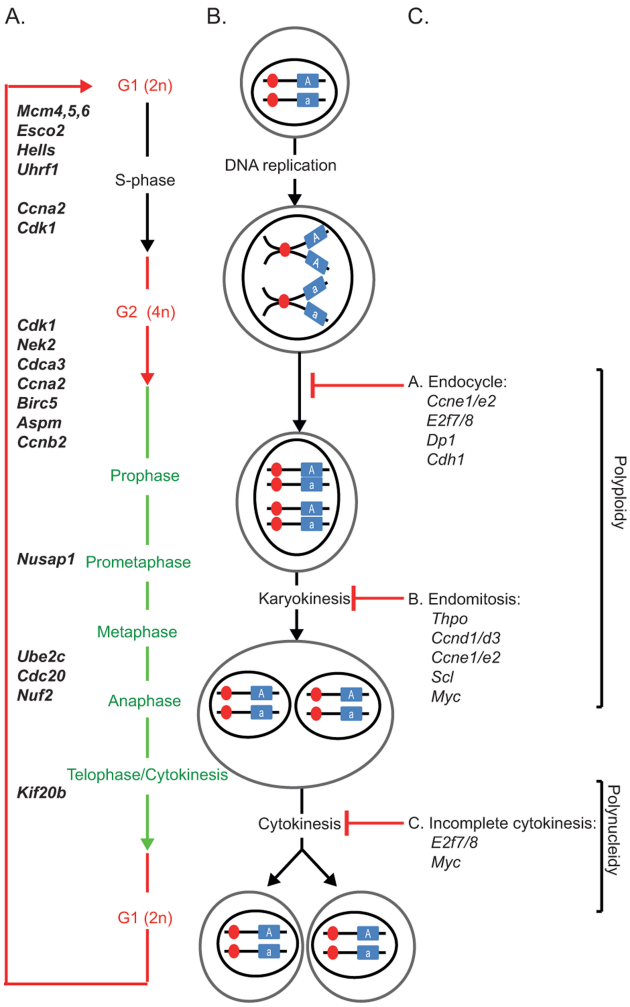

FIG. 6. Mouse and human CAR/PXR mediate similar PB-induced upregulation of genes driving both entry into S-phase and progression to cytokinesis in wild-type and humanized mice. (A) Schematic representation of different cell cycle stages; a selection of genes which are both reported to regulate the different stages and are significantly upregulated upon PB treatment between day 1 and day 3 of PB treatment is shown. (B) Schematic representation of DNA content along the different cell cycle stages. (C) Polyploidy or polynucleidy can result from incomplete cytokinesis, endocycle, or endomitosis that are likely regulated by genes reviewed in Pandit *et al.* (2013). Genes in bold-italic in panel (A) are differentially upregulated between day 1 and day 7 upon PB treatment in wild-type animals and CAR^{h}-PXR^{h}, but not in CAR^{KO}-PXR^{KO}, consistent with a PB-induced hepatic proliferative transcriptional response. No PB-induced transcriptional responses were observed in either wild-type animals or CAR^{h}-PXR^{h} for genes associated with polyploidy or polynucleidy in panel (C).

and binding domains share unusually low amino acid sequence conservation for nuclear receptor orthologs (Moore *et al.*, 2002) which may account for differences in respective receptor activities upon PB exposure, as the ligand-binding domain indirectly affects other receptor functions such as dimerization, ligand binding, interaction with heat shock proteins, nuclear localization, and transactivation functions. Indeed, structural differences in human CAR versus mouse CAR translated into structure-activity differences in the ability of different ligands to activate or deactivate CAR in comparative assays (Moore *et al.*, 2000). In contrast to the species differences observed for some CAR ligands, our data suggest that the indirect activation of both mouse and human CAR by PB leads to very similar hepatic xenobiotic and proliferative transcriptional responses in a C57BL/6 mouse genetic background. Although we did not find any compelling evidence for PB-mediated transcriptional responses that were specific to humanized CAR/PXR mice, multiple genes (e.g., *Meig1*, *Prom1*, *Knck1*, *Cyp2b13*) were observed to exhibit quantitatively distinct transcriptional responses to PB in CAR^{h}-PXR^{h} versus wild-type mice and these did not appear to be driven by differences in liver PB exposure. These quantitative differences are thus likely to reflect target gene regulatory sequence and/or chromatin structural differences in mouse CAR- and human CAR-mediated DNA binding and recruitment of mouse liver transcriptional coregulatory proteins. Consistent with this notion, a precedent for species-specific transcription factor interactions with their target genes has recently been reported (Soccio *et al.*, 2011).

Our data have important implications for the utility of humanized CAR/PXR mouse models in human nongenotoxic carcinogenesis risk assessment. Although the CAR^{h}-PXR^{h} mice used in our studies express human CAR splice variants 1–3 in a ratio that closely resembles human liver CAR expression profile and is unchanged by exposure to PB (Ross *et al.*, 2010), one caveat of this model is that human CAR and PXR receptors function in the context of mouse target gene regulatory elements and chromatin structure. Previous work has shown that PB induces extensive changes in DNA and histone modification patterns across the regulatory regions of CAR target genes in mouse liver (Lempiainen *et al.*, 2011; Phillips and Goodman, 2009; Thomson *et al.*, 2012, 2013), and species-specific differences in these epigenetic perturbations may also play an important role in determining susceptibility to PB-mediated hepatocarcinogenesis. Further species differences in CAR signaling might also be conferred via the recently described PB-EGFR signaling pathway interactions (Mutoh *et al.*, 2013).

Based on a weight of evidence human relevance framework concept focusing on mode of action and key events, PB-induced rodent nongenotoxic hepatocarcinogenesis is not considered to be a relevant mechanism for humans (Holsapple *et al.*, 2006) and there is no evidence of a specific role of PB in human liver cancer risk based on epidemiological data in epileptics (IARC, 2001; Lamminpaa *et al.*, 2002; La Vecchia and Negri, 2013; Whysner *et al.*, 1996). The fact that CAR activation is a key event for PB-induced rodent liver tumor formation (Elcombe *et al.*, 2014) suggests that the use of humanized CAR mouse models may more closely reflect human transcriptional responses and should be more predictive of human risk. However, our data suggest that humanized nuclear receptor mice

may not be a simple model for extrapolating the risk of rodent tumor findings to humans. Understanding and using these models will require the careful integration of quantitative exposure-response relationships with the temporal and spatial dynamics of human nuclear receptor expression, mechanism of modulation by coactivators and further evaluation of the relevance of heterologous mouse-human gene regulatory protein interactions.

SUPPLEMENTARY DATA

Supplementary data are available online at http://toxsci.oxfordjournals.org/.

FUNDING

Innovative Medicine Initiative Joint Undertaking (IMI JU) (115001; MARCAR project; URL: http://www.imi-marcar.eu/);Novartis and Bayer Pharma; Cancer Research UK Programme (C4639/A12330 to C.R.W); MRC UK Core Programme (632RMERA1892 to R.R.M.).

ACKNOWLEDGMENTS

All IMI-MARCAR consortium partners had a role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We would like to thank Ulrich Längle and Andreas Mahl for *in vivo* study support. V.D., A.M., S-D.C., H.H., F.P., D.T., P.C., O.G., P.M., R.T., and J.G.M. are full time employees of Novartis Pharma AG. H.L. and A.V. are recipients of Novartis Institutes for Biomedical Research Postdoctoral Fellowships. N.S. is a full time employee of Taconic Artemis. H.E-Z. is a full time employee of Bayer. C.R.E. is a full time employee of CXR Biosciences.

REFERENCES

Auerbach, S. S., Stoner, M. A., Su, S., and Omiecinski, C. J. (2005). Retinoid X receptor-alpha-dependent transactivation by a naturally occurring structural variant of human constitutive androstane receptor (NR1I3). *Mol. Pharmacol.* **68,** 1239–1253.

Aydinlik, H., Nguyen, T. D., Moennikes, O., Buchmann, A., and Schwarz, M. (2001). Selective pressure during tumor promotion by phenobarbital leads to clonal outgrowth of beta-catenin-mutated mouse liver tumors. *Oncogene* **20,** 7812–7816.

Becker, F. F. (1982). Morphological classification of mouse liver tumors based on biological characteristics. *Cancer Res.* **42,** 3918–3923.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B (Methodological)* **57,** 289–300.

Bohm, N., and Noltemeyer, N. (1981). Excessive reversible phenobarbital induced nuclear DNA-polyploidization in the growing mouse liver. *Histochemistry* **72,** 63–74.

Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19,** 185–193.

Bursch, W., Grasl-Kraupp, B., Wastl, U., Hufnagl, K., Chabicovsky, M., Taper, H., and Schulte-Hermann, R. (2004). Role of apoptosis for mouse liver growth regulation and tumor promotion: Comparative analysis of mice with high (C3H/He) and low (C57Bl/6J) cancer susceptibility. *Toxicol. Lett.* **149,** 25–35.

Elcombe, C. R., Peffer, R. C., Wolf, D. C., Bailey, J., Bars, R., Bell, D., Cattley, R. C., Ferguson, S. S., Geter, D., Goetz, A. et al. (2014). Mode of action and human relevance analysis for nuclear receptor-mediated liver toxicity: A case study with phenobarbital as a model constitutive androstane receptor (CAR) activator. *Crit. Rev. Toxicol.* **44,** 64–82.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J. et al. (2004) Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.* **5,** R80.

Gentric, G., Celton-Morizur, S., and Desdouets, C. (2012a). Polyploidy and liver proliferation. *Clin. Res. Hepatol. Gastroenterol.* **36,** 29–34.

Gentric, G., Desdouets, C., and Celton-Morizur, S. (2012b) Hepatocytes polyploidization and cell cycle control in liver physiopathology. *Int. J. Hepatol.* **2012,** 282430.

Gonzales, A. J., Christensen, J. G., Preston, R. J., Goldsworthy, T. L., Tlsty, T. D., and Fox, T. R. (1998). Attenuation of G1 checkpoint function by the non-genotoxic carcinogen phenobarbital. *Carcinogenesis* **19,** 1173–1183.

Hasmall, S. C., and Roberts, R. A. (1999). The perturbation of apoptosis and mitosis by drugs and xenobiotics. *Pharmacol. Ther.* **82,** 63–70.

Holsapple, M. P., Pitot, H. C., Cohen, S. M., Boobis, A. R., Klaunig, J. E., Pastoor, T., Dellarco, V. L., and Dragan, Y. P. (2006). Mode of action in relevance of rodent liver tumors to human cancer risk. *Toxicol. Sci.* **89,** 51–56.

Huang, W., Zhang, J., Washington, M., Liu, J., Parant, J. M., Lozano, G., and Moore, D. D. (2005). Xenobiotic stress induces hepatomegaly and liver tumors via the nuclear receptor constitutive androstane receptor. *Mol. Endocrinol.* **19,** 1646–1653.

In monographs, I. (Ed.)ppIARC (2001). Phenobarbital and its sodium salt. In: *Some thyrotropic agents*, pp. 161–286.

Kauffmann, A., Gentleman, R., and Huber, W. (2009). arrayQualityMetrics—A bioconductor package for quality assessment of microarray data. *Bioinformatics* **25,** 415–416.

Kawamoto, T., Sueyoshi, T., Zelko, I., Moore, R., Washburn, K., and Negishi, M. (1999). Phenobarbital-responsive nuclear translocation of the receptor CAR in induction of the CYP2B gene. *Mol. Cell. Biol.* **19,** 6318–6322.

Lamminpaa, A., Pukkala, E., Teppo, L., and Neuvonen, P. J. (2002). Cancer incidence among patients using antiepileptic drugs: A long-term follow-up of 28,000 patients. *Eur. J. Clin. Pharmacol.* **58,** 137–141.

Lehmann, J. M., McKee, D. D., Watson, M. A., Willson, T. M., Moore, J. T., and Kliewer, S. A. (1998). The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions. *J. Clin. Invest.* **102,** 1016–1023.

Lempiainen, H., Couttet, P., Bolognani, F., Muller, A., Dubost, V., Luisier, R., Espinola Adel, R., Vitry, V., Unterberger, E. B., Thomson, J. P. et al. (2013). Identification of Dlk1-Dio3 imprinted gene cluster noncoding RNAs as novel candidate biomarkers for liver tumor promotion. *Toxicol. Sci.* **131,** 375–386.

Lempiainen, H., Muller, A., Brasa, S., Teo, S. S., Roloff, T. C., Morawiec, L., Zamurovic, N., Vicart, A., Funhoff, E., Couttet, P. et al. (2011) Phenobarbital mediates an epigenetic switch at the constitutive androstane receptor (CAR) target gene Cyp2b10 in the liver of B6C3F1 mice. *PLoS One* **6,** e18216.

Moore, L. B., Maglich, J. M., McKee, D. D., Wisely, B., Willson, T. M., Kliewer, S. A., Lambert, M. H., and Moore, J. T. (2002). Pregnane X receptor (PXR), constitutive androstane receptor (CAR), and benzoate X receptor (BXR) define three pharmacologically distinct classes of nuclear receptors. *Mol. Endocrinol.* **16,** 977–986.

Moore, L. B., Parks, D. J., Jones, S. A., Bledsoe, R. K., Consler, T. G., Stimmel, J. B., Goodwin, B., Liddle, C., Blanchard, S. G., Willson, T. M. et al. (2000). Orphan nuclear receptors constitutive androstane receptor and pregnane X receptor share xenobiotic and steroid ligands. *J. Biol. Chem.* **275,** 15122–15127.

Mutoh, S., Sobhany, M., Moore, R., Perera, L., Pedersen, L., Sueyoshi, T., and Negishi, M. (2013) Phenobarbital indirectly activates the constitutive active androstane receptor (CAR) by inhibition of epidermal growth factor receptor signaling. *Sci. Signal.* **6,** ra31.

Nims, R. W., Sinclair, P. R., Sinclair, J. F., Dragnev, K. H., Jones, C. R., Mellini, D. W., Thomas, P. E., and Lubet, R. A. (1993). Dose-response relationships for the induction of P450 2B by 1,4-bis[2-(3,5-dichloropyridyloxy)]benzne (TCPOBOP) in rat and cultured rat hepatocytes. *Xenobiotica* **23,** 1411–1426.

Pandit, S. K., Westendorp, B., and de Bruin, A. (2013). Physiological significance of polyploidization in mammalian cells. *Trends Cell Biol.* **23,** 556–566.

Pandit, S. K., Westendorp, B., Nantasanti, S., van Liere, E., Tooten, P. C., Cornelissen, P. W., Toussaint, M. J., Lamers, W. H., and de Bruin, A. (2012). E2F8 is essential for polyploidization in mammalian cells. *Nat. Cell Biol.* **14,** 1181–1191.

Parzefall, W., Erber, E., Sedivy, R., and Schulte-Hermann, R. (1991). Testing for induction of DNA synthesis in human hepatocyte primary cultures by rat liver tumor promoters. *Cancer Res.* **51,** 1143–1147.

Pfaffl, M. W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29,** e45.

Phillips, J. M., Burgoon, L. D., and Goodman, J. I. (2009a). The constitutive active/androstane receptor facilitates unique phenobarbital-induced expression changes of genes involved in key pathways in precancerous liver and liver tumors. *Toxicol. Sci.* **110,** 319–333.

Phillips, J. M., Burgoon, L. D., and Goodman, J. I. (2009b). Phenobarbital elicits unique, early changes in the expression of hepatic genes that affect critical pathways in tumor-prone B6C3F1 mice. *Toxicol. Sci.* **109,** 193–205.

Phillips, J. M., and Goodman, J. I. (2009). Multiple genes exhibit phenobarbital-induced constitutive active/androstane receptor-mediated DNA methylation changes during liver tumorigenesis and in liver tumors. *Toxicol. Sci.* **108,** 273–289.

Pirttiaho, H. I., Sotaniemi, E. A., Pelkonen, R. O., and Pitkanen, U. (1982). Hepatic blood flow and drug metabolism in patients on enzyme-inducing anticonvulsants. *Eur. J. Clin. Pharmacol.* **22,** 441–445.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing..

Rignall, B., Braeuning, A., Buchmann, A., and Schwarz, M. (2011). Tumor formation in liver of conditional beta-catenin-deficient mice exposed to a diethylnitrosamine/phenobarbital tumor promotion regimen. *Carcinogenesis* **32,** 52–57.

Ross, J., Plummer, S. M., Rode, A., Scheer, N., Bower, C. C., Vogel, O., Henderson, C. J., Wolf, C. R., and Elcombe, C. R. (2010). Human constitutive androstane receptor (CAR) and pregnane X receptor (PXR) support the hypertrophic but not the hyperplastic response to the murine nongenotoxic hepatocarcinogens phenobarbital and chlordane in vivo. *Toxicol. Sci.* **116,** 452–466.

Saito, K., Morre, R., and Negishi, M. (2013). Nuclear receptor CAR specifically activates the two-pore K+ channel Kcnk1 gene in male mouse livers, which attenuates phenobarbital-induced hepatic hyperplasia.. *Toxicol Sci.* **132,** 151–161.

Scheer, N., and Wolf, C. R. (2013). Genetically humanized mouse models of drug metabolizing enzymes and transporters and their applications. *Xenobiotica* **44,** 96–108.

Scheer, N., Ross, J., Kapelyukh, Y., Rode, A., and Wolf, C. R. (2010). In vivo responses of the human and murine pregnane X receptor to dexamethasone in mice. *Drug Metab. Dispos.* **38,** 1046–1053.

Scheer, N., Ross, J., Rode, A., Zevnik, B., Niehaves, S., Faust, N., and Wolf, C. R. (2008). A novel panel of mouse models to evaluate the role of human pregnane X receptor and constitutive androstane receptor in drug response. *J. Clin. Invest.* **118,** 3228–3239.

Scheer, N., Snaith, M., Wolf, C. R., and Seibler, J. (2013). Generation and utility of genetically humanized mouse models. *Drug Discov. Today* **18,** 1200–1211.

Shizu, R., Benoki, S., Numakura, Y., Kodama, S., Miyata, M., Yamazoe, Y., and Yoshinari, K. (2013) Xenobiotic-induced hepatocyte proliferation associated with constitutive active/androstane receptor (CAR) or peroxisome proliferator-activated receptor alpha (PPARalpha) is enhanced by pregnane X receptor (PXR) activation in mice. *PLoS One* **8,** e61802.

Soccio, R. E., Tuteja, G., Everett, L. J., Li, Z., Lazar, M. A., and Kaestner, K. H. (2011). Species-specific strategies underlying conserved functions of metabolic transcription factors. *Mol. Endocrinol.* **25,** 694–706.

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., Kono, T., Shioda, T., and Hochedlinger, K. (2010). Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **465,** 175–181.

Styles, J. A. (1993). Measurement of ploidy and cell proliferation in the rodent liver. *Environ. Health Perspect.* **101**(Suppl. 5)**,** 67–71.

Thomson, J. P., Hunter, J. M., Lempiainen, H., Muller, A., Terranova, R., Moggs, J. G., and Meehan, R. R. (2013). Dynamic changes in 5-hydroxymethylation signatures underpin early and late events in drug exposed liver. *Nucleic Acids Res.* **41,** 5639–5654.

Thomson, J. P., Lempiainen, H., Hackett, J. A., Nestor, C. E., Muller, A., Bolognani, F., Oakeley, E. J., Schubeler, D., Terranova, R., Reinhardt, D. et al. (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. *Genome Biol.* **13,** R93.

Tolson, A. H., and Wang, H. (2010). Regulation of drug-metabolizing enzymes by xenobiotic receptors: PXR and CAR. *Adv. Drug Deliv. Rev.* **62,** 1238–1249.

Ueda, A., Hamadeh, H. K., Webb, H. K., Yamamoto, Y., Sueyoshi, T., Afshari, C. A., Lehmann, J. M., and Negishi, M. (2002). Diverse roles of the nuclear orphan receptor CAR in regulating hepatic genes in response to phenobarbital. *Mol. Pharmacol.* **61,** 1–6.

La Vecchia, C., and Negri, E. (2013). A review of epidemiological data on epilepsy, phenobarbital, and risk of liver cancer. *Eur. J. Cancer Prev.* **23,** 1–7.

Whysner, J., Ross, P. M., and Williams, G. M. (1996). Phenobarbital mechanistic data and risk assessment: Enzyme induction, enhanced cell proliferation, and tumor promotion. *Pharmacol. Ther.* **71,** 153–191.

Yamamoto, Y., Moore, R., Goldsworthy, T. L., Negishi, M., and Maronpot, R. R. (2004). The orphan nuclear receptor constitutive active/androstane receptor is essential for liver tumor promotion by phenobarbital in mice. *Cancer Res.* **64,** 7197–7200.

Supplementary Material

Luisier Raphaëlle[1], Lempiaïnen Harri[1], Scherbichler Nina[1], Braeuning Albert[2], Geissler Miriam[2], Dubost Valérie[1], Müller Arne[1], Scheer Nico[3], Chibout Salah-Dine[1], Hara Hisanori[4], Frank Picard[4], Theil Diethilde[1], Couttet Philippe[1], Vitobello Antonio[1], Grenet Olivier[1], Grasl-Kraup Bettina[5], Ellinger-Ziegelbauer Heidrun [6], Thomson John P.[7], Meehan Richard R.[7], Elcombe Clifford R.[8], Henderson Colin[9], Wolf C. Roland[9], Michael Schwarz[2], Moulin Pierre[1],Terranova Rémi [1,*], Moggs Jonathan G.[1,*]

[1]Preclinical Safety, Novartis Institutes for Biomedical Research,
CH-4057 Basel, Switzerland

[2]Department of Toxicology, Institute of Experimental and Clinical Pharmacology and Toxicology,
University of Tübingen, 72074 Tübingen, Germany

[3]TaconicArtemis, Neurather Ring 1, Köln 51063, Germany

[4]DMPK, Novartis Institutes for Biomedical Research,
CH-4057 Basel, Switzerland

[5]Institute of Cancer Research, University of Vienna, Austria

[6]Department of Investigational Toxicology,Bayer Pharma AG, Germany

[7]Breakthrough Breast Cancer Research Unit and Division of Pathology, University of Edinburgh,
Western General Hospital, Edinburgh EH4 2XU, United Kingdom

[8]CXR Biosciences Ltd, 2 James Lindsay Place, Dundee DD1 5JJ, UK

[9]Medical Research Institute, University of Dundee,
Ninewells Hospital & Medical School, Dundee, UK

[*]Corresponding authors: Terranova Rémi (remi.terranova@novartis.com) & Moggs Jonathan G. (jonathan.moggs@novartis.com),
CHBS WKL135, Klybeckstrasse 141, CH-4057 Basel, Switzerland

January 31, 2014

1 Supplementary protocol

1.1 Affymetrix mRNA and microRNA GeneChip processing

Processing of GeneChip® experiments was conducted as recommended by the manufacturer of the GeneChip® system (Affymetrix, Santa Clara, CA). For tissue samples, double stranded cDNA was synthesized with a starting amount of 0.1 μg total RNA. For RNA reverse transcription, the GeneChip® 3' IVT Express Labeling Assay (lot ID 0904012, Affymetrix) was used in the presence of a T7-(dT)24 DNA oligonucleotide primer (Affymetrix). The cDNA was then transcribed *in vitro* in the presence of biotinylated ribonucleotides to form biotin-labelled amplified RNA (aRNA). The labelled aRNA was then purified and quantified by UV spectrophotometry at 260 nm and fragmented. 10 μg of fragmented biotinylated aRNA were hybridized for approximately 16 hrs at 45 °C and 48 °C to the GeneChip® Mouse430_2 arrays and GeneChip® miRNA2.0 arrays respectively. The arrays were then washed and stained with the GeneChip® Hybridization Wash and Stain kit (Affymetrix). The washing and staining steps were performed with GeneChip® Fluidics Workstation 450 (Affymetrix). Arrays were then scanned using a solid-state laser scanner (GeneArray® Scanner 3000 combined with the GeneChip® autoloader, Affymetrix). The Affymetrix GeneChip® Operating Software (GCOS) was used to generate the primary and secondary raw data files. The scanned images from miRNA were converted into numerical values of the signal intensity (Signal) and into categorical expression level measurement (Absolute Call) using the Affymetrix AGCC software.

1.2 In situ hybridization (ISH) and Immunohistochemistry (IHC)

Template for Meg3 riboprobe synthesis was generated by RT-PCR on RNA from mouse brain using self-priming oligonucleotide primers flanked in 5' with SP6- and T3-promoter recognition sequences (forward primer: SP6CTCTTCTC CATCGAACGGCT, reverse primer T3-AACAATAAAGAACTTGAAGAGGTTTTGAT, amplicon size: 537 bp). The purified PCR product was transcribed using T3-RNA polymerase (anti-sense) and

SP6-RNA polymerase (sense) at 37 °C for 2 hrs using dNTP containing Digoxigenin-UTP according to the manufacturer recommendations (Roche Diagnostics, Schweiz AG, Rotkreuz, Switzerland). The quality and quantity of the riboprobe was evaluated using the 2100 Bioanalyzer. ISH was performed using the fully automated instrument Ventana Discovery Ultra® (Roche Diagnostics). All chemicals were also provided by Roche Diagnostics. Briefly, formalin fixed paraffin embedded sections were de-paraffinized and rehydrated under solvent-free conditions (EZprep solution). Pretreatment steps were done with the RiboMap™ kit following the manufacturers instructions followed by cell conditioning (demasking) performed by heat retrieval cycles in RiboCC solution using option mild followed by a complementary enzymatic digestion (Protease 3 for 16 minutes at 37 °C). Hybridization was performed adding to each slide 200 μl of RiboHybe solution containing 10 ng of DIG-riboprobe and incubating at 70 °C for 6 hrs. After hybridization section were washed 3 times at 70 °C for 8 min on stringency conditions (2.0 x SSC). DIG-label probe detection was performed using an Alkaline Phosphatase-conjugated Sheep anti-Digoxigenin antibody (Roche Diagnostics) diluted 1/500 in antibody diluent. Antibody incubation was carried out for 30 min at 37 °C followed by chromogenic detection using BlueMap™ Kit with a substrate incubation time of 4hrs. Counterstaining using ISH nuclear fast red was performed for 2 min. Sections were mounted in Glycerol-gelatin mounting medium (Sigma-Aldrich Chemie GmbH, Buchs, Switzerland) and post-mounted using Pertex™. For double staining with Glutamine synthetase (GS), the rabbit anti-GS antibody from Sigma (catalog number G2781) was used at a dilution of 1/20'000 in antibody diluent for 3 hrs and was applied just after the alkaline phosphatase conjugated sheep anti-digoxigenin. The detection step was immediately done using a biotin conjugated donkey anti-rabbit antibody (dilution 1/500 in antibody diluent and incubation time 16 min) followed by application of the DABmap™ kit according to the provider recommendations. The chromogenic detection for the DIG-labeled probe using the BlueMap Kit was done at the end. IHC for Ki67 was performed using the fully automated instrument Ventana Discovery® (Roche Diagnostics). All chemicals were also provided by Roche Diagnostics. Formalin fixed paraffin embedded sections were de-paraffinized and rehydrated under solvent-free conditions (EZprep solution) followed by antigen retrieval (demasking) performed by heat retrieval cycles in a Tris-EDTA based buffer (CC1 solution, option standard). Subsequently slides were blocked using 1x Casein solution in PBS (BioFX laboratories Inc, Catalog number PBSC-0100-5x) and endogenous avidin/biotin activity was quenched for 4 min. Some 100 μl of a rabbit anti-Ki67 from NeoMarker (catalog number RM-9106S) diluted at 1/200 in antibody diluent were added on slides and incubated for 3 hrs at room temperature. A short post-fixation (glutaraldehyde at 0.05%) was done before applying a biotin conjugated donkey anti-rabbit at 1/500 for 16 min (Jackson Immunoresearch Inc.). Detection was performed with a streptavidin-biotin peroxidase detection system DABMap® Kit following the manufacturer recommendations. Slides were counter stained with Hematoxylin and bluing reagent, dehydrated and mounted using Pertex™ (Biosystems Switzerland AG, Nunningen, Switzerland).

2 Supplementary figures

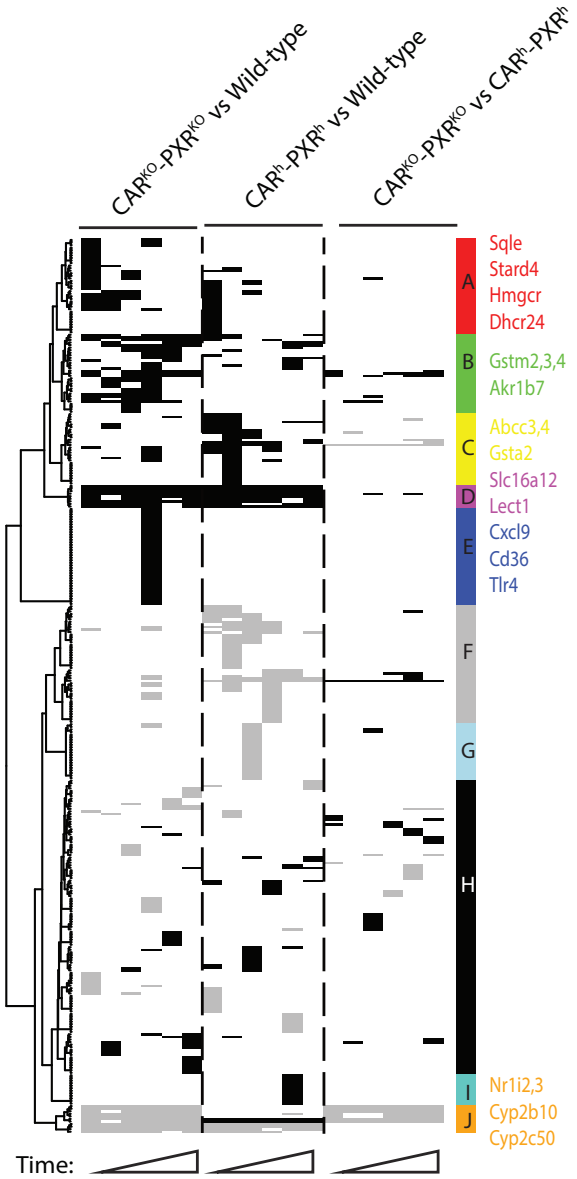

Figure S1: Summary of differential gene expression due to strain effect: comparison of gene expression at corresponding time-point of control CAR^{KO}-PXR^{KO} versus control wild-type, control CAR^{h}-PXR^{h} versus control wild-type, and control CAR^{KO}-PXR^{KO} versus control CAR^{h}-PXR^{h} mouse livers. Black dots = genes significantly up-regulated, grey dots = genes significantly down-regulated and white dots = no significant change. A gene is considered significantly up-regulated if $|\log_2 \mathrm{FC}| > 0.53$ (corresponding to FC>1.5 or FC<0.69) and B.H. (Benjamini and Hochberg) corrected P-Value < 0.01. Genes are clustered hierarchically by (1) computing Euclidean distance between genes from decision matrix and (2) applying Ward clustering algorithm. Detailed gene list can be consulted in Supplementary material (2).

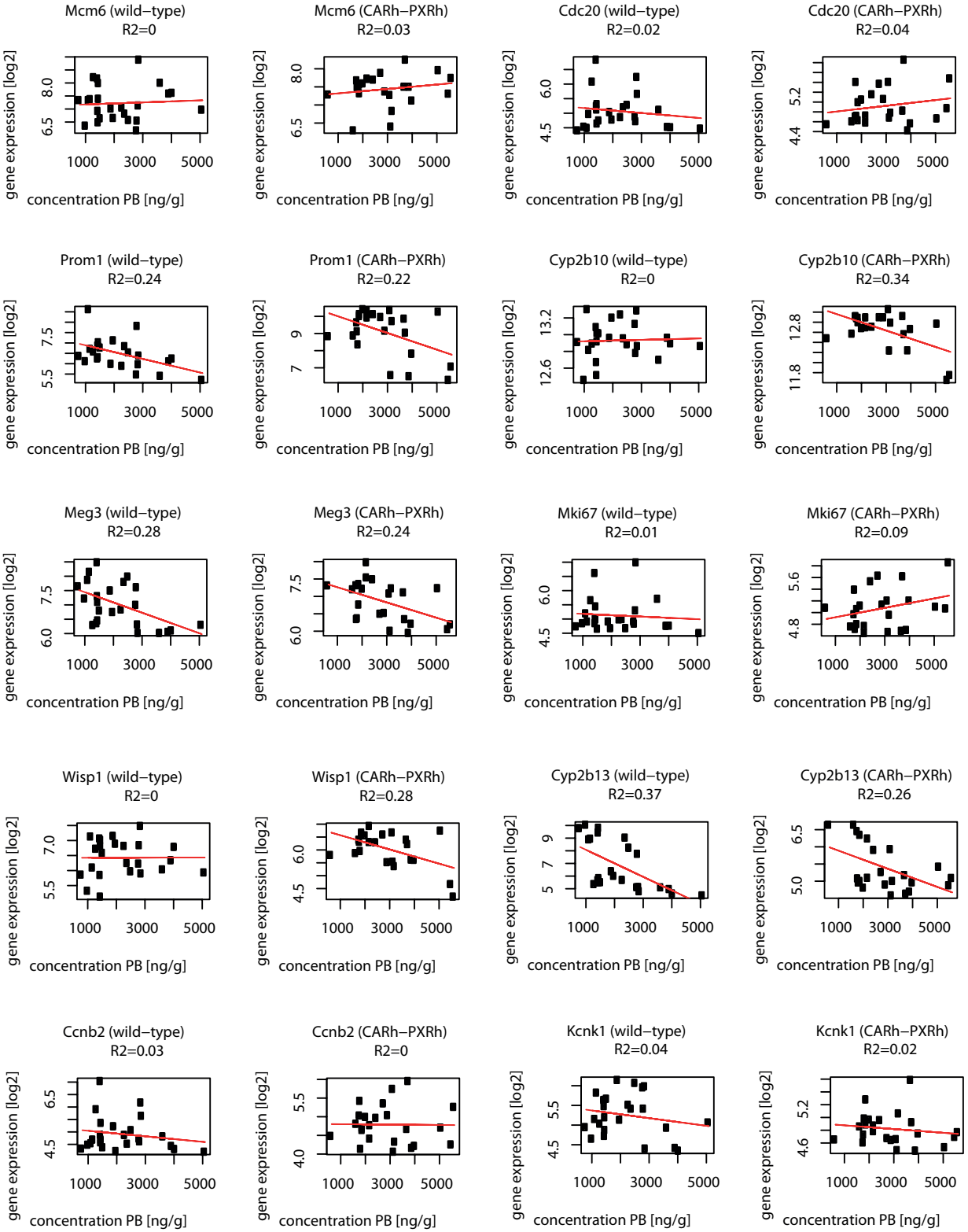

Figure S2: Linear modeling between gene expression and PB liver exposure for the genes of interest. Results do not support significant effect on gene induction upon changes in PB liver exposure. Whilst as for *Cyp2b13*, linear modeling suggests anti-correlation between PB exposure and gene expression, this indeed results from differential expression over time: as PB liver exposure decreases over time, *Cyp2b13* gene expression increases over time.

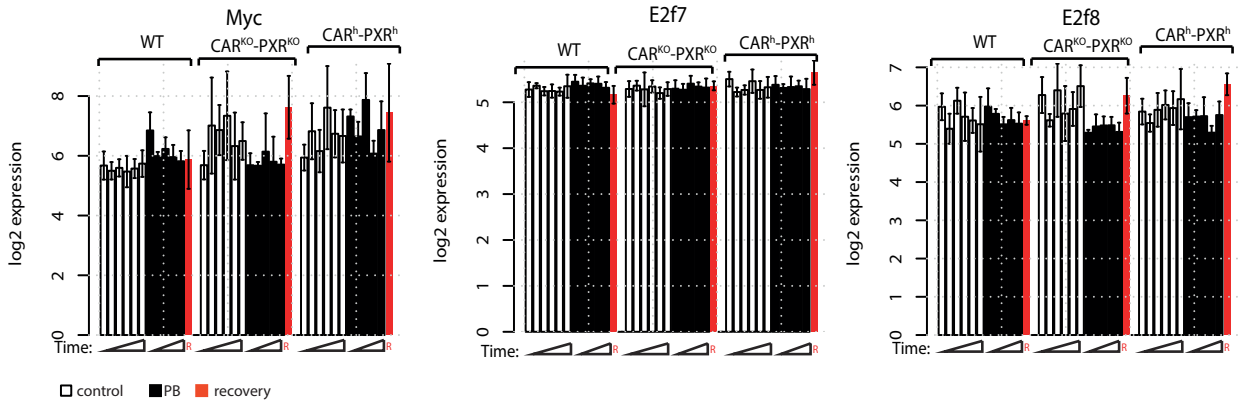

Figure S3: Expression of *E2F7*, *E2F8* and *Myc* upon PB treatment in different strains. Expression ($\log_2$) in control (open bars), treated (black bars) and recovery time-point (red bars) male animals is given as mean±SD (n=3-5 animals per group).

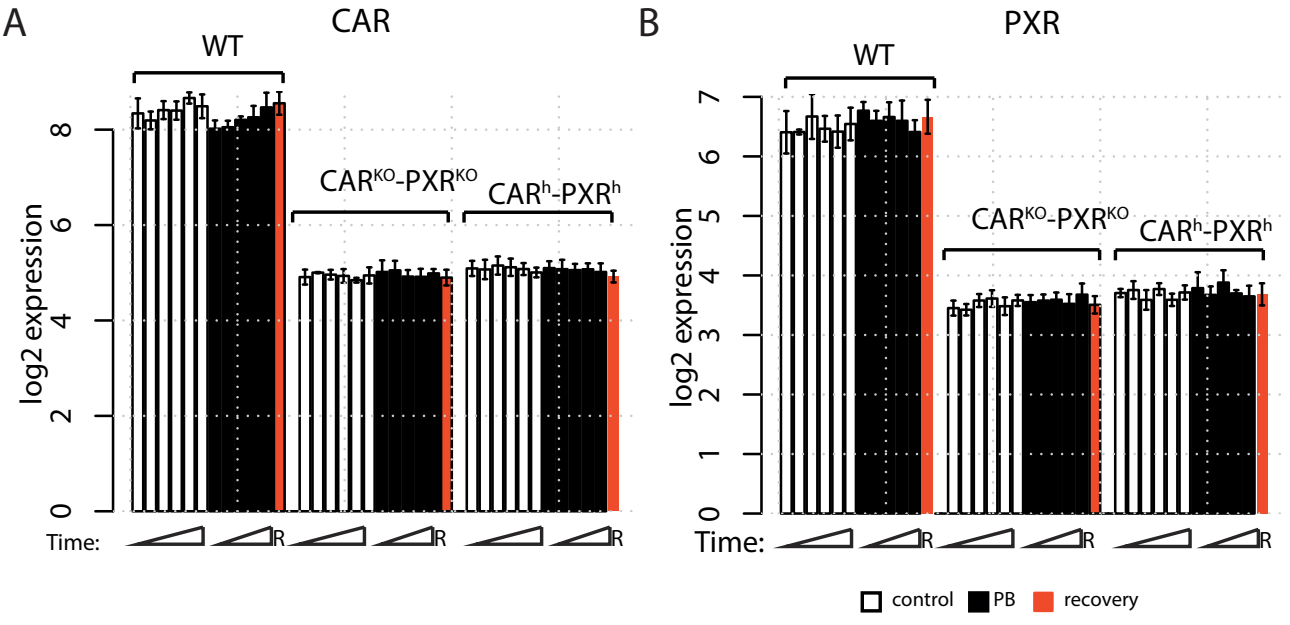

Figure S4: Expression of CAR and PXR in different strains. Log-expression lower than 5.0 is considered as background signal. Expression ($\log_2$) in control (open bars), treated (black bars) and recovery time-point (red bars) male animals is given as mean±SD (n=3-5 animals per group).

3 Supplementary tables

Liver						
	Wild-type		**CAR^{KO}-PXR^{KO}**		**CAR^{h}-PXR^{h}**	
Time[day]	**mean [ng/g]**	**sd**	**mean [ng/g]**	**sd**	**mean [ng/g]**	**sd**
1	39220	7946	58220	11362	44420	10680
7	16680	6796	50400	12400	26800	7556
14	20260	5533	48520	20617	29160	14417
28	21620	6637	57800	15399	25840	5757
91	11874	2892	36960	9567	14483	5465
Plasma						
	Wild-type		**CAR^{KO}-PXR^{KO}**		**CAR^{h}-PXR^{h}**	
Time[day]	**mean [ng/mL]**	**sd**	**mean [ng/mL]**	**sd**	**mean [ng/mL]**	**sd**
1	52400	4900	58700	9370	54100	8890
7	19400	3660	71900	10900	28900	5420
14	21100	3860	75300	14500	31500	7730
28	19700	3590	82700	8860	32300	5180
91	11100	3220	52600	6250	17000	7230

Table S1: Plasma and liver PB exposure as measured using LC-MS/MS at each time-point in treated animals (n=3-5 animals per group).

Gene Symbol	Gene Description	Gene Function	Day1				Day 7			
			WT		CAR^h-PXR^h		WT		CAR^h-PXR^h	
			$\log_2$ FC	B.H. P-Val	$\log_2$ FC	B.H. P-Val	$\log_2$ FC	B.H. P-Val	$\log_2$ FC	B.H. P-Val
Cenpk	Centromere protein K Gene	Component of the nucleosome distal complex which is involved in assembly of kinetochore proteins, mitotic progression and chromosome segregation.	0.8	1.3E-05	0.2	5.7E-01	0.7	1.0E-03	0.6	5.2E-03
Ect2	Cct2 oncogene Gene	Binds highly specifically to RhoA, RhoC and Rac proteins, but does not appear to catalyze guanine nucleotide exchange.	0.9	2.0E-02	1.0	7.1E-03	1.8	8.3E-08	1.3	9.5E-05
Esco2	Establishment of cohesion 1 homolog 2	Required for the establishment of sister chromatid cohesion and couple the processes of cohesion and DNA replication to ensure that only sister chromatids become paired together.	0.7	6.1E-02	0.5	2.8E-01	1.4	1.7E-05	1.3	1.3E-04
Fign11	Fidgetin-like 1	May regulate osteoblast proliferation and differentiation	1.2	1.0E-10	0.6	5.0E-03	1.0	6.4E-06	0.6	5.5E-03
Glt25d2	Glycosyltransferase 25 domain containing 2	Has a beta-galactosyltransferase activity; transfers beta-galactose to hydroxylysine residues on collagen	0.9	1.4E-09	0.8	6.3E-07	0.6	6.5E-04	0.7	4.1E-05
Hells	Helicase, lymphoid specific	Required for de novo or maintenance DNA methylation. May play a role in formation and organization of heterochromatin, implying a functional role in the regulation of transcription and mitosis	1.4	2.2E-08	0.6	8.6E-02	1.0	2.7E-03	0.9	5.4E-03
Mcm4	Minichromosome maintenance deficient 4 homolog	Involved in the control of DNA replication	0.7	6.2E-05	0.6	6.3E-03	0.7	9.0E-04	0.8	1.0E-04
Mcm5	Minichromosome maintenance deficient 5, cell division cycle 46		1.4	5.4E-07	1.0	1.5E-03	1.1	2.6E-03	1.0	1.4E-03
Mcm6	Minichromosome maintenance deficient 6	Binds to chromatin during G1 and detaches from it during S phase, implying that it allows the chromatin to replicate	1.3	3.6E-05	1.1	9.5E-04	1.2	6.0E-04	1.5	1.6E-06
Mki67	Antigen identified by monoclonal antibody Ki 67		0.6	9.3E-02	0.6	9.5E-02	1.1	7.4E-04	0.8	1.4E-02
Ncapg2	Non-SMC condensin II complex, subunit G2	Regulatory subunit of the condensin-2 complex	0.8	2.4E-04	0.4	1.7E-01	0.8	9.0E-04	0.5	8.1E-02
Papss2	3'-phosphoadenosine 5'-phosphosulfate synthase	Bifunctional enzyme with both ATP sulfurylase and APS kinase activity, which mediates two steps in the sulfate activation pathway.	1.0	9.2E-14	0.6	1.5E-05	0.6	4.1E-04	0.1	7.9E-01
Pbk	PDZ binding kinase	Phosphorylates MAP kinase p38. Seems to be active only in mitosis.	0.7	2.1E-03	0.3	5.3E-01	1.0	8.5E-05	0.8	2.4E-03
Uhrf1	Ubiquitin-like, containing PHD and RING finger domains, 1	Putative E3 ubiquitin-protein ligase. May participate in methylation-dependent transcriptional regulation. Important for G1/S transition. May be involved in DNA repair and chromosomal stability.	0.8	2.0E-05	0.5	1.2E-02	0.8	1.4E-04	0.7	1.3E-03

Table S2: PB-mediated differentially expressed genes (from Day 1 until Day 7) functionally enriched in DNA replication. Subset of cluster C of Figure 2 and group (1) of Figure 5. Gene function obtained from **STRING 9.05 - Known and Predicted Protein-Protein Interactions**.

Gene Symbol	Gene Description	Gene Function	Day1				Day 7			
			WT		CAR^{h}-PXR^{h}		WT		CAR^{h}-PXR^{h}	
			$\log_2$ FC	B.H. P-Val	$\log_2$ FC	B.H. P-Val	$\log_2$ FC	B.H. P-Val	$\log_2$ FC	B.H. P-Val
Arhgap11a	Rho GTPase activating protein 11A		0.2	4.4E-01	0.2	5.4E-01	0.6	1.1E-03	0.5	8.0E-03
Aspm	Abnormal spindle-like, microcephaly associated	Probable role in mitotic spindle regulation and coordination of mitotic processes.	0.3	2.2E-01	0.1	7.1E-01	0.6	2.2E-03	0.4	6.8E-02
Birc5	Baculoviral IAP repeat-containing 5	Component of the chromosomal passenger complex (CPC), a complex that acts as a key regulator of mitosis.	0.4	3.7E-01	0.4	4.3E-01	1.4	1.0E-07	0.9	3.9E-03
Ccna2	Cyclin A2	Essential for the control of the cell cycle at the G1/S (start) and the G2/M (mitosis) transitions	0.4	1.0E-01	0.4	1.1E-01	1.0	1.5E-05	0.9	7.2E-05
Ccnb2	Cyclin B2	Essential for the control of the cell cycle at the G2/M (mitosis) transition	0.2	8.0E-01	0.2	7.8E-01	1.9	2.8E-10	1.3	6.3E-06
Cdc20	Cell division cycle 20 homolog	Required for full ubiquitin ligase activity of the anaphase promoting complex/cyclosome (APC/C)	0.3	3.7E-01	0.4	1.9E-01	1.6	9.4E-15	1.0	1.1E-06
Cdca3	Cell division cycle associated 3	F-box-like protein which is required for entry into mitosis.	0.2	7.2E-01	0.4	1.4E-01	1.1	2.4E-08	0.9	2.5E-06
Cdk1	Cyclin-dependent kinase 1	Required in higher cells for entry into S-phase and mitosis. p34 is a component of the kinase complex that phosphorylates the repetitive C-terminus of RNA polymerase II	0.3	3.4E-01	0.3	3.4E-01	0.7	5.7E-03	0.8	9.6E-04
Cdkn3	Cyclin-dependent kinase inhibitor 3		0.2	2.4E-01	0.2	3.8E-01	0.7	5.0E-08	0.4	3.2E-04
Ckap2	Cytoskeleton associated protein 2	Possesses microtubule stabilizing properties. Involved in regulating aneuploidy, cell cycling, and cell death in a p53- dependent manner	0.3	1.2E-01	0.2	6.3E-01	1.0	2.0E-10	0.8	7.3E-07
Dtl	Denticleless homolog	Seems to be necessary to ensure proper cell cycle regulation of DNA replication.	0.5	2.0E-02	0.4	3.3E-02	0.6	3.1E-03	0.5	7.1E-03
Gsta2	Glutathione S-transferase, alpha 2	Conjugation of reduced glutathione to a wide number of exogenous and endogenous hydrophobic electrophiles	0.3	4.6E-01	0.3	4.3E-01	0.7	8.8E-03	0.8	6.2E-04
Gstt1	Glutathione S-transferase, theta 1	Conjugation of reduced glutathione to a wide number of exogenous and endogenous hydrophobic electrophiles.	0.4	7.3E-06	0.2	1.2E-01	0.6	4.9E-10	0.2	7.3E-02
Hmmr	Hyaluronan mediated motility receptor	Involved in cell motility.	0.3	7.7E-02	0.3	3.7E-02	0.8	7.0E-11	0.7	2.8E-09
Kif20b	Kinesin family member 20B	Plus-end-directed motor enzyme that is required for completion of cytokinesis	0.4	6.6E-02	0.1	7.9E-01	0.7	1.1E-03	0.5	3.7E-02
Nek2	Never in mitosis gene a-related expressed kinase 2	Protein kinase involved in mitotic regulation. May have a role at the G2-M transition.	0.2	3.5E-01	0.0	1.0E+00	0.8	1.1E-05	0.7	1.0E-04
Nuf2	NUF2, NDC80 kinetochore complex component	Acts as a component of the essential kinetochore-associated NDC80 complex, which is required for chromosome segregation and spindle checkpoint activity.	0.3	2.2E-01	0.4	1.0E-01	0.6	1.9E-03	0.4	1.0E-01
Nusap1	Nucleolar and spindle associated protein 1	Microtubule-associated protein with the capacity to bundle and stabilize microtubules.	0.3	2.6E-01	0.3	3.6E-01	0.7	8.7E-03	0.2	5.4E-01
Pard3b	Par-3 partitioning defective 3 homolog B	Putative adapter protein involved in asymmetrical cell division and cell polarization processes	0.5	1.5E-04	0.3	1.7E-01	0.6	5.1E-05	0.3	1.3E-01
Rapgef4	Rap guanine nucleotide exchange factor	Guanine nucleotide exchange factor (GEF) for RAP1A, RAP1B and RAP2A small GTPases that is activated by binding cAMP.	0.5	3.7E-01	-0.1	9.4E-01	1.2	2.5E-03	0.1	9.4E-01
Stmn1	Stathmin 1	Involved in the regulation of the microtubule filament system by destabilizing microtubules.	0.2	7.9E-01	0.1	9.4E-01	0.8	2.3E-03	0.8	1.5E-03
Ube2c	Ubiquitin-conjugating enzyme E2C	Acts as an essential factor of the anaphase promoting complex/cyclosome (APC/C), a cell cycle-regulated ubiquitin ligase that controls progression through mitosis.	0.4	3.4E-01	0.2	6.6E-01	0.9	4.4E-04	0.7	7.4E-03
Zwilch	Zwilch, kinetochore associated, homolog	Essential component of the mitotic checkpoint, which prevents cells from prematurely exiting mitosis.	0.5	3.2E-02	0.4	1.8E-01	0.9	2.5E-05	0.6	7.3E-03

Table S3: PB-mediated differentially expressed genes (around Day 7) functionally enriched in mitosis. Subset of cluster C of Figure 2 and group (2) of Figure 5. Gene function obtained from **STRING 9.05 - Known and Predicted Protein-Protein Interactions**.

Chapter 5

Discussion and concluding remarks

In this research project we have applied novel bioinformatic approaches to comprehensive gene expression data from various *in vivo* studies leading to (i) identification of early regulators of PB-induced liver tumorigenesis and (ii) assessment of relevance of CARPXR humanized mouse model in testing receptor-dependent mechanisms underlying liver tissue molecular responses to NGC. In the following we discuss the main findings, their impact in drug safety assessment, and future direction.

5.1 New DNA-binding regulators of early PB-mediated liver tumorigenesis

A better understanding of the regulatory mechanisms underlying long-term effects of non-genotoxic carcinogens and more particularly identification of the TFs that regulate PB-mediated long-term transcriptional changes down-stream of CAR and β-catenin is of great importance for drug development and safety assessment of the carcinogenic potential of compounds with similar mode of action. In this study we have adapted a robust *ab initio* probabilistic algorithm that models gene expression dynamics in terms of predicted cis-regulatory sites to comprehensive toxicogenomic data from *in vivo* experiments to propose new DNA-binding regulators involved in regulation of early PB mediated liver tumor promotion.

5.1.1 Summary of major findings

Collectively these analyses propose new and testable regulatory mechanisms underlying three key aspects PB-mediated tumor promotion that are 1) PB-mediated DNA replication in hepatocytes (**Figure 5.1b**), 2) PB-mediated xenobiotic response (**Figure 5.1-c**), and 3) PB-mediated tumor-prone environment establishment (**Figure 5.1-d**). Additionally these analyses identified several candidate regulators of 1) liver tumorigenesis (**Figure 5.1-e**), 2) early PB-mediated kinetics of transcriptional response down-stream of CAR signaling and 3) liver context candidate regulators down-stream of β-catenin signaling pathway.

5.1.1.1 Regulators underlying early PB-mediated kinetics of transcriptional response

The transcriptional response mediated by long-term PB treatment has been previously reported to be complex and non-linearly dependent on time (123). Although CAR has been shown to initiate most of PB-mediated transcriptional response, its activity cannot be the sole regulator of this response. Using singular value decomposition applied to motif activity matrix obtained from MARA, we quantified and characterized TFs activity variations associated with specific biological processes i.e. constant xenobiotic

response, transient PB-mediated mitogenic response and adaptive xenobiotic response, and identified regulators underlying these biological pathways. Thus we propose new candidate CAR-down-stream regulators of early PB-mediated kinetics of early transcriptional response.

5.1.1.2 New liver-specific β-catenin down-stream regulators

β-catenin has been shown to physically interact with diverse nuclear transcription factors (169). As β-catenin role in PB-mediated liver tumor promotion is more likely to arise from co-operation with alternative TF activated by PB rather than direct activation by the compound, a list of liver-specific β-catenin down-stream regulators would be very informative. In this study we systematically interrogated 190 motifs for β-catenin down-stream regulation by adapting MARA algorithm to genomic data from β-catenin KO samples and proposed 30 candidate co-factors of β-catenin in the liver. Importantly the method successfully identified TCF/LEF cofactor binding site motif which is the best characterized co-factor of Wnt/β-catenin signaling pathway (203; 259).

5.1.1.3 E2F as a positive regulator of the PB-mediated mitogenic response at both the early and tumor stages

An important aspect of PB-mediated tumor promotion is the ability of PB to induce transient mitogenic response and cause liver neoplasia on chronic administration. Previous studies suggest that different populations of hepatocytes are sensitive to proliferative induction according to stimuli such as chemical exposure and reduction in liver mass (39; 40; 41; 42). However the exact mechanisms responsible for the exit from the quiescent state and the re-entry into the cell cycle remain largely unknown (see (39) for review). We here hypothesized that motifs similarly dysregulated during early PB-mediated mitogenic response and in promoted tumors specifically were strong candidates for HC cell proliferation culminating in liver cancer.

Our analysis revealed an increase in E2F motif activity in PB-mediated proliferative tissues (early transient peak of proliferation and tumor stage), but none in non-promoted tumors, suggesting distinct regulatory programs of cell proliferation between these two tumor types. Numerous studies report indeed central role of distinct E2Fs family members in HCC (260; 261). However we are the first to our knowledge to show specific modulation of the motif in promoted liver tumors. Furthermore our analysis revealed that E2F motif activity is negatively modulated upon β-catenin KO, suggesting a positive interaction between these two pathways. Given the constitutive activation of β-catenin in promoted tumors, amplified interaction between these two regulators may lead to aberrant E2F activity.

Whilst from MARA result, we cannot determine which of the E2F family members is responsible for these changes in activity, significant correlation between *E2f1, E2f2* and *E2f8* gene expression and motif activity suggests that these are likely to bind E2F motif. Moreover our analysis revealed that both *myc* and *E2f8* are 1) predicted target genes of E2F and 2) significantly up-regulated in promoted tumors only. Interestingly both c-Myc (262; 263) and E2F8 (56) are key regulators of hepatocyte polyploidization, a mechanism which contributes to increase in liver metabolic load. It is therefore tempting to speculate that PB treatment induces E2F8 and c-Myc activation, potentially through EGFR signaling inhibition, leading to increase in polyploid cells rather that increase in cell number through incomplete cytokinesis. Indeed the measurement of PB-induced mouse liver proliferation is confounded by changes in a heterogenous liver parenchymal cell population that include both mononuclear and binucleated cells (264; 188). Thus, increases in either Ki67 mRNA expression, Ki-67/PCNA IHC staining or BrdU incorporation previously reported in PB-treated hepatocytes and described as a proliferative response may indeed reflect an in-

creased in DNA content of a polyploidy hepatocyte rather than DNA replication associated with complete cell division cycles and hyperplasia. Further histopathologic and flow cytometry endpoints are required to rigorously differentiate between increased ploidy in hypertrophic hepatocytes versus hepatocellular proliferation, and in order to assess differential activity of E2F family upon PB treatment.

5.1.1.4 ZFP161 as transcriptional repressor involved in the PB-mediated mitogenic response at both the early and tumor stages

Our analysis revealed a negative modulation of motif bound by ZFP161 (also known as ZF5) upon PB treatment contributing to the early transient mitogenic. In addition, ZFP161 targets are down-regulated in promoted tumors, but not in non-promoted tumors. ZFP161 has been shown to be preferentially active in differentiated tissues with little mitotic activity (265), where it is likely to act as a transcriptional repressor, and it is thought to directly repress *myc* (266; 267). However predicted targets for ZFP161 in this study are enriched in transcriptional repressors for cell proliferation (i.e. *Mxi1* and *Klf10*), which are themselves down-regulated in proliferative tissues whilst *myc* is not predicted as target and is up-regulated in these tissues. We therefore hypothesize that ZFP161 participates in the PB-mediated regulation of quiescent hepatocytes G0-G1 transition at both the early and tumor stages by repressing negative regulators of cell cycle and this mechanism is specific to hepatocytes. Importantly while very few studies were done on this TF, we are the first to propose a regulatory role for ZFP161 in liver tissue that is further supported by the relatively high expression level of ZFP161 in liver samples.

5.1.1.5 ESR1 repression and creation of a tumor prone environment

PB-mediated tumorigenesis involves dynamic changes in tissue composition that progressively create a tumor-prone environment, resulting from the liver adaptive response to chronic stress. Our analysis identified ESR1 as a factor progressively down-regulated upon PB chronic exposure at early stage and specifically down-regulated in promoted tumors, making this TF a strong candidate regulator for this process. Furthermore analysis of the motif activity changes upon β-catenin KO revealed a negative interaction between ESR1 and β-catenin down-stream signaling supporting the hypothesis of progressive inhibition of ESR1 signaling by β-catenin constitutive activation. ESR1 motif activity and TF expression correlate significantly in β-catenin and tumor studies only, which coincides with samples where β-catenin activity is predicted to change. These results support the hypothesis of a negative ESR1 transcriptional regulation by β-catenin. Of note physical interaction between β-catenin/TCF-4 and ESR1 in other physiological contexts was reported elsewhere (268; 269). PB-mediated inhibition of EGFR may be an additional mechanism of ESR1 inhibition, given that estrogen-independent EGFR-dependent activation of ESR1 was shown elsewhere (270; 271).

While ESR1 tumor suppressor activity is supported by various studies (107; 128; 272; 273; 274) we are the first to propose a PB-mediated progressive suppression of ESR1 activity upon chronic exposure as one of the mechanisms underlying PB-mediated liver tumor promotion. The tumor suppressor role of ESR1 evidenced in previous studies was shown to result partly from estrogen-mediated inhibition of IL-6 expression in Kupffer cells that in turn affect hepatocyte proliferation (128) suggesting key role for ESR1 in hepatocyte communication with non-parenchymal cells. Moreover results from few studies performed in both males and females rat over more than 2 years suggest that PB promotional effect is similar in males and females, whereas female mice live longer under PB-treatment (135; 136; 137). We think that the time-delay between male and female liver outcome upon PB treatment results in part by higher basal ESR1 activity in females compared to males leading female mice to less sensitivity to PB-mediated decrease in ESR1 activity.

According to our predictions β-catenin negatively regulates ESR1 activity, at least in males. Direct interaction between ESR1 and TCF4 was indeed previously shown (268) and ESR1-liver specific activity may depend on such an interaction. Upon β-catenin constitutive activation, as it is the case in the Ctnnb1 mutated cells, β-catenin may interact with most available TCF4, leading to fewer free TCF4 and resulting in decrease ESR1-TCF4 complex formation eventually leading to decrease in activity.

5.1.2 Future work and experimental follow-up

As all these hypotheses are the results of computational predictions, further experimental follow-up is necessary to ascertain the relevance of these regulators in PB-mediated tumor promotion. Because the pathogenesis of PB-mediated tumor promotion is long (35 weeks) and results also from communication between hepatocytes and non-parenchymal cells, long-term effects of PB exposure are difficult to test using *in vitro* culture of hepatocytes that rapidly differentiate on plastic and loose expression of CYPs. However short term effects such as PB-mediated transient proliferative response can be in principle reproduced in freshly isolated primary hepatocytes and thus regulators involved in early PB-mediated proliferative response should in principle be testable *in vitro*. In the following we propose experiments to either validate or further investigate our hypotheses, some of them having being initiated in the laboratory of Carcinogenesis and Epigenetics (PCS, Novartis, Basel).

5.1.2.1 Characterization of proliferative index and ploidy

Given the predicted modulation of E2F activity specifically in promoted tumors, the recently demonstrated role for E2F8 and Myc in hepatocyte ploidy, and the increase in hepatocyte ploidy upon PB exposure, we speculate that a PB-mediated coordinated change in ploidy rather than a increase in proliferative index is responsible for the predicted increase in motif activity. In order to characterize proliferative index and ploidy in both promoted and non-promoted tumors, we have initiated immunostaining of paraffin-embedded promoted and non-promoted tumors with Ki67, a marker of proliferation, and Feulgen, that allows to quantify ploidy (275; 276). We expect similar proliferative index between the two tumor types but differences in ploidy.

5.1.2.2 Assessment of changes in TFs activity in promoted tumors specifically

TFs activity is regulated at various levels such as transcriptional level (these changes were already tested by differential gene expression analysis), post-translational modification of the protein, changes in cellular localization, changes in expression of co-factors and DNA accessibility of target genes (through demethylation of promoters for example). Immunohistochemistry staining of promoted and non-promoted tumors samples together with surrounding normal tissue would allow (i) to assess that the proteins are present in tissues of interest, (ii) to identify the cells that express the TFs (please note that gene expression data are very sensitive and thus tissue contamination with other cells than hepatocytes can lead to significant signal possibly responsible for motif activity change), and (iii) to test for differential TFs activity as per changes in cellular localization (either nuclear accumulation, or switch from membranous to cytoplasmic). Furthermore changes in motif activity that either result from interaction with newly available co-factors and for changes DNA accessibility of the targets could be tested using ChIP assays that would inspect for changes in binding of TFs between the different conditions. Please note however that these experiments necessitate highly specific antibodies. Antibody testing was initiated in our laboratory and very few antibodies so far yielded high enough specificity.

5.1.2.3 Assessment of PB-mediated modulation of TF activity

In order to test for PB-mediated changes in TFs activity, we propose to perform a motif-reporter activity assay in freshly isolated hepatocytes under various conditions such as with and without PB, with and without CAR, with and without β-catenin, in order to genetically measure PB-mediated TFs signaling and investigate dependency on different regulators. Furthermore, in order to assess which of the E2F family member is responsible for PB-mediated change in activity motif at early stage, combination with siRNA treatment would allow to test each of the family members separately.

5.1.2.4 Study of biochemical protein interactions

In order to test for biochemical protein interactions between β-catenin and the various identified regulators down-stream of β-catenin signaling, we propose a co-immunoprecipitation (Co-IP) experiment using cellular extracts from freshly isolated mouse hepatocytes.

5.1.3 Discussion about current approach

Given that PB mediates changes in gene expression primarily through inhibition of EGFR (186), it is likely that many predicted changes in TFs activities result from post-translational modifications rather than changes in TFs gene expression. MARA infers regulatory activities from the behavior of predicted targets thus this algorithm allows for prediction of differential activity which may be due to post-translational modifications, changes in cellular localization, or interactions with co-factors. Therefore we think that the current approach is more powerful in identifying candidate TFs for tumor promotion than classic methods such as analysis of differentially expressed genes or enrichment of motifs in differentially expressed genes. Moreover MARA provides with a list of predicted targets for each motif allowing to interrogate for biological pathways regulated by the TFs of interest.

However this algorithm is also limited by the TFBS data-base: in mammals, sequence-specificities are available for only about 350 of the about 1,500 TFs. While CAR is the most important regulator of PB-mediated tumor promotion and the most characterized, prediction for CAR would have been a good way of validating the method. However there is to our knowledge no reliable motif for CAR/PXR neither in JASPAR(277; 86) nor in TRANSFAC (87). An important step would therefore be the creation of ChIP-seq data in treated liver for CAR. However until now we have not been able to find a reliable CAR antibody.

Another limitation of the method is that MARA focuses solely on predicted TFBSs in proximal promoters, ignoring the effects of distal enhancers. Expanding the region would lead to increase in false TFBS eventually leading to noisy signal. However previous studies in liver for example have shown CYPs gene regulation by Ahr in enhancers regions and this cannot be captured by the current model. Moreover motifs mode of action cannot be distinguished and therefore TFs which act as much as positive as negative regulators will generate a zero signal as per weighted average of every predicted promoter expression.

5.1.4 Relevance to safety assessment of drugs in development

As reviewed earlier, identification of NGC requires 2-years *in vivo* experiments in rodents. Indeed one of the major mechanistic difference between a genotoxic and a non-genotoxic carcinogen is the time-scale of pathogenesis. As tissue exposure to a sufficient dose of genotoxic agent results in direct DNA mutagenesis and DNA repair response, the carcinogenic effects of NGC results from progressive changes in tissue homeostasis and interplay between several cellular subtypes. By applying a robust *ab initio*

probabilistic modeling to toxicogenomics data we here propose a new set of hypotheses that can be readily tested. Compared to the 1,500 known TFs in rodents that could potentially be involved in PB-mediated tumor promotion, our results are expected to significantly speed up the process of understanding the regulatory mechanisms underlying the pathogenesis. Moreover we are confident that the identification of early regulators of NGC can be valuable early biomarkers for NGC and improve safety assessment of compounds with similar mode of action. Finally a better understanding of the regulatory mechanisms of PB-mediated tumorigenesis provide insights into PB MOA that can help to understand human relevance of these models. In conclusion although this study is mainly based on the analysis of publicly available transcriptome that require additional experimental validations, the resulting hypotheses highlight novel potential early mechanisms and pathways for liver tumor promotion, providing new opportunities for early assessment of the carcinogenic potential of therapeutic compounds.

5.1.5 Remaining open questions and hypothese

Our analysis propose robust new regulators for transcriptional response underlying different biological processes such as PB-mediated hepatocyte proliferation and the repression of regulators with demonstrated tumor suppressor activity that are down-stream of β-catenin. However the following two important questions regarding PB-mediated tumor promotion remain unsolved: 1) how does PB promote outgrowth of β-catenin activated cells that under physiological condition are non-tumorigenic and 2) how does PB repress the outgrowth of H-ras activated cells that are, in absence of PB, highly tumorigenic.

As reviewed earlier, in the absence of PB treatment, 30% of DEN-initiated tumors harbors activating mutation in H-ras, leading to activation of MAPK signaling, whereas under PB-treatment 80% of the tumors are mutated in β-catenin. Furthermore high similarity between gene expression patterns of H-ras and β-catenin tumors with gene expression patterns from periportal and perivenous hepatocytes respectively (278) suggests distinct and mutually exclusive regulatory programs between these two tumor types. Interestingly while β-catenin mutation alone is insufficient for hepatocarcinogenesis and H-ras mutation alone rapidly causes large cell dysplasia in the hepatocytes, simultaneous induction of H-ras and β-catenin mutations leads to 100% HCC development supporting positive and cooperative interaction between these two pathways (279). This cooperative interaction between H-ras and β-catenin pathway has indeed been proposed in a recent review (280), where cooperation between the two pathways was hypothesized to trigger midzonal proliferation after 2/3 hepatectomy as a result from an extension in the β-catenin and the H-ras activation territories leading to overlapping activities (282). This idea is supported by the finding that while neither H-ras mutation nor β-catenin activation led to urothelial cell carcinoma (UCC) (within 12 months), mice carrying both mutations rapidly developed UCC (281). We consequently propose the following mechanisms for PB-mediated selection of β-catenin mutated cells and PB-mediated repression of H-ras mutated cells.

PB-mediated negative selection for β-catenin activated cells Because PB treatment selects for β-catenin activated cells, it is reasonable to think that PB negatively selects for cells that are able to escape β-catenin degradation through ubiquitination by accelerating β-catenin degradation, leading to an overall decrease in β-catenin activity. Because glutamine synthetase is not affected by PB treatment, we think that PB-mediated degradation of β-catenin is effective solely in regions where basal β-catenin activity is low i.e. everywhere else than in perivenous area.

PB-mediated repression of H-ras mutated cells Given that simultaneous induction of H-ras and β-catenin mutations leads to 100% HCC development, and considering our hypothesis for PB-mediated β-

catenin inactivation true, the impairment for H-ras cell proliferation may be due to complete inactivation of β-catenin in periportal cells upon PB treatment. Of note the characterization of β-catenin nuclear content using immunohistochemistry in hepatocytes failed in demonstrating preferential nuclear amount in pervineous relative to periportal hepatocytes.

PB-mediated outgrowth of β-catenin activated cells First it is noteworthy that PB-mediated outgrowth of β-catenin cells is a process that requires at least 24 weeks of PB exposure although pre-existing mutated cells are likely to be present in tissue from the beginning. This strongly suggests that progressive changes in hepatocyte regulatory programs and/or non-parenchymal cellular compartment are necessary to promote for β-catenin activated cells. Given that β-catenin activated cells under normal condition is not sufficient for HCC development whereas simultaneous induction of H-ras and β-catenin mutations leads to 100% HCC development, we hypothesize that PB mediates progressive H-ras activation. Once the level of H-ras activation is high enough in cells bearing β-catenin mutations, cells can start to grow. Of note alternative pathway may be progressively activated by PB (or non-parenchymal cells invading the tissue) that cooperates with β-catenin.

In conclusion we think that PB treatment leads to 1) accelerated degradation of β-catenin, prohibiting H-ras cell proliferation that require basal β-catenin activity and 2) progressive up-regulation of a complementary pathway (potentially H-ras) that cooperate with β-catenin to trigger hepatocyte proliferation. As sufficient β-catenin activity is only found in mutated cells, PB eventually promotes tumors with β-catenin mutation.

5.2 Human relevance of humanized CARPXR mouse model

Based on epidemiological data in epileptics, long-term barbiturate treatment is not associated with increased incidence of liver tumors in humans (214; 215; 283). Furthermore human hepatocytes are resistant to the ability of PB to increase cell proliferation (212; 213) and inhibit apoptosis (213) whilst prolonged treatment with PB does increase liver size in humans (284; 211). These differences in species biochemistry and pathophysiology have raised doubts regarding the appropriateness of extrapolating some rodent tumor findings to humans (9). Thus a better understanding of MOA of such compounds is believed to help addressing the relevance of rodent assays to human risk assessment (10; 9).

As reviewed earlier CAR is as a key TF that is required for liver tumor formation elicited upon prolonged PB treatment in mice. Consequently an important question towards addressing translatability of rodent model to humans is to test for distinct physiological roles and behaviors of human and murine CAR upon PB exposure. Similar cellular mechanisms of PB-induced mouse and human CAR nuclear translocation and activation have been demonstrated that partly arise from dephosphorylation of a threonine residue in the very well conserved DNA binding domain (DBD) of the receptor (184). However human and mouse CAR have unusual divergent ligand-binding domains (LBDs) that may explain possible different roles of CAR in these species (285; 286). Indeed LBDs mediate ligand binding, heterodimerization with co-factors such as RXR (285), interaction with heat shock proteins, nuclear localization, and transactivation functions as reviewed in (287) and the LBD of CAR has been suggested to bind coactivators such as SRC-1 (179) or TIF2/GRIP1 (288) in the absence of ligands to exhibit its constitutive activity (116). Therefore, despite similar DNA binding domaines (DBD), divergent LBD between human and mouse can lead to different physiological roles. In this study, we have used a humanized CAR/PXR mouse model to examine potential species differences in receptor-dependent mechanisms underlying liver tissue molecular responses to PB.

5.2.1 Resume of major findings

First, wild-type and CAR^{h}-PXR^{h} mouse livers exhibited temporally and quantitatively similar transcriptional responses during 91 days of phenobarbital exposure including the sustained induction of the xenobiotic response gene *Cyp2b10*, the Wnt signalling inhibitor *Wisp1* and non-coding RNA biomarkers from the *Dlk1-Dio3* locus. Second, transient induction of DNA replication (*Mcm6, Esco2, Uhrf1*) and mitotic genes (*Ccnb2; Cdc20; Cdk1*) and the proliferation-related nuclear antigen *Mki67* were observed with peak expression occurring between 1 and 7 days PB exposure in both wild-type and humanized mice. All these transcriptional responses were absent in CAR^{KO}-PXR^{KO} mouse livers and largely reversible in wild-type and CAR^{h}-PXR^{h} mouse livers following 91 days of PB exposure and a subsequent 4 week recovery period. Furthermore, PB-mediated up-regulation of the non-coding RNA *Meg3*, that has recently been implicated in cellular reprogramming, exhibited a similar dose response and peri-venous hepatocyte-specific localization in both wild-type and CAR^{h}-PXR^{h} mice. Thus mouse livers co-expressing human CAR and PXR can support xenobiotic and proliferative transcriptional responses following exposure to PB in mouse context .

5.2.2 Limitations of humanized model

Despite that most transcriptional and phenotypic responses upon PB exposure were very similar in humanized and wild-type models, some differences were observed between the two strains both under physiological condition and under PB treatment that reminds the existence of clear species differences between human and mouse nuclear receptors, like, for example, divergent LBDs leading to potential alternative cofactors depending on species. Therefore the utility of humanized model towards addressing species differences is questionable and conclusions have to be drawn carefully.

As reviewed earlier PB-induced mouse and human CAR nuclear translocation and activation are similar due to similar DBD domains. However additional cellular mechanisms are required for CAR DNA binding such as heterodimerization with a variety of cofactors depending on gene and tissue. Indeed one caveat of this model is that human nuclear receptors function in the context of mouse gene regulatory elements and proteins i.e. different co-factors. We can not exclude that human and mouse CAR/PXR ability to trigger hepatocyte proliferation upon PB exposure in mouse context may result from heterodimerization with a cofactor specific to mouse context. The absence of such cofactor from human context would be one reason for human resistance of PB to increase cell proliferation (212; 213). Thus a co-IP experiment that would provide with a species-specific list of CAR co-factors would help addressing this issue. Alternatively ChIP experiments on human and mouse primary hepatocytes treated with PB would enable the direct comparison of PB-induced target genes bound by CAR. Finally experiments with "murinized" CAR-PXR human hepatocytes treated with PB would enable to assess importance of mouse context in PB-mediated human CAR-PXR induced mouse hepatocyte proliferation.

5.2.3 Implications for drug safety assessment

Based on a weight of evidence human relevance framework concept focusing on mode of action and key events, PB-induced rodent non-genotoxic hepatocarcinogenesis is not considered to be a relevant mechanism for humans (9) and there is no evidence of a specific role of PB in human liver cancer risk based on epidemiological data in epileptics (214; 215). Our data suggest that humanized nuclear receptor mice may not be a simple model for extrapolating the risk of rodent tumor findings to humans and will require the careful integration of quantitative exposure-response relationships with the temporal and spatial dynamics of human nuclear receptor expression, mechanism of modulation by coactivators and

relevance of heterologous mouse-human gene regulatory protein interactions.

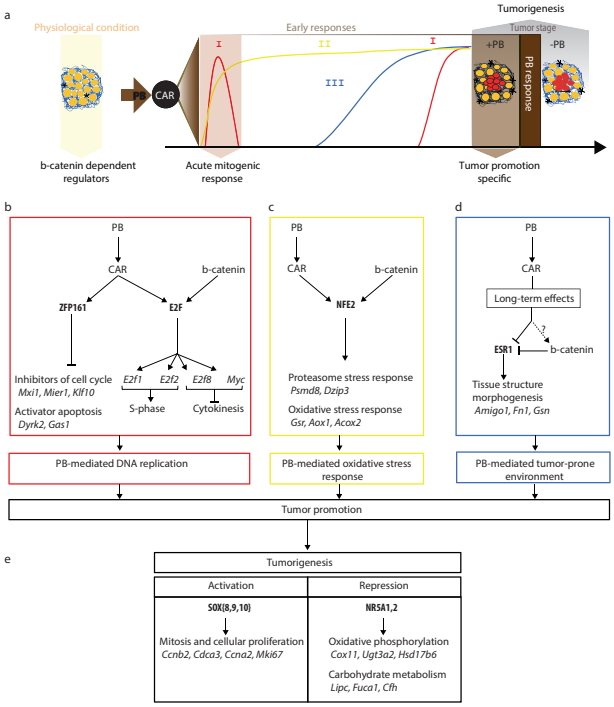

Figure 5.1: Schematical representation of PB-mediated tumor promotion, as it emerges from the study. **(a)** Illustration of the PB-mediated tumor promotion process and the aspects elucidated by the 4 experimental studies that we analyze. Knock-out of β-catenin identifies regulators downstream of β-catenin in physiological conditions (yellow arrow). Ours and previous analyses suggest that all regulatory effects of PB-treatment are downstream of CAR activation (brown arrow and black circle). Our motif activity and SVD analysis of the early kinetic time course identified three key biological processes induced by PB-treatment: A transient mitogenic response, which is also associated with a late resurgence of proliferation (I, red), a sustained xenobiotic response (II, yellow), and a late response which is likely involved in establishing a tumor prone environment (III, blue). Comparison of promoted and non-promoted tumors identifies motifs dysregulated in all tumors, and in promoted tumors only (grey arrows). **(b-d)** Summary of the key regulators of liver tumor promotion organized according to biological (colored boxes matching the colors of processes I, II, and III in panel a) with arrows indicating regulatory interactions between regulators and on selected target genes. **(b)** E2F and ZFP161 regulate PB-mediated hepatocyte proliferation at the early and promoted tumor stage. E2F is down-stream of β-catenin signaling and likely induces both DNA replication, via up-regulation of *E2f1,2*, and aborted cytokinesis via up-regulation of *E2f8* and *c-myc*. ZFP161 is likely involved in the G0-G1 transition via transcriptional repression of transcriptional repressors of cell growth and cell cycle. **(c)** NFE2, downstream of β-catenin as well, is involved in the sustained xenobiotic response, upregulating proteasome activity and the oxidative stress response. **(d)** PB-mediated suppression of ESR1 activity underlies development of a tumor-prone environment, most likely through repression of tissue morphogenesis. β-catenin signaling represses ESR1. **(e)** Key regulators involved in tumorigenesis, i.e. disregulated in both promoted and non-promoted tumors. Increased SOX{8,9,10} activity likely regulates hepatocyte mitosis and proliferation via up-regulation of cyclins. Decrease in NR5A1,2 activity is detected after 3 months of PB treatment and maintained in tumor samples, and therefore a good early indicator of hepatocyte loss of function associated with tumorigenesis.

Chapter 6

Concluding remark

In the course of this research project we have brought new insights and directly testable hypotheses into regulatory mechanisms of drug-induced non-genotoxic carcinogenesis. Furthermore human relevance of rodent models in drug toxicity assessment was discussed in the context of humanized mouse model. We are confident that this work provides with a good example of how applying sophisticated mathematical modeling to toxicogenomic data can significantly speed up the process of biomarker identification and provide with a better understanding of mechanisms and pathways underlying drug-induced toxicity.

Bibliography

[1] Silva Lima, B. and Van der Laan, J. W. (2000 Oct) Mechanisms of nongenotoxic carcinogenesis and assessment of the human hazard. *Regul Toxicol Pharmacol*, **32**, 135–143.

[2] LeBaron, M. J., Rasoulpour, R. J., Klapacz, J., Ellis-Hutchings, R. G., Hollnagel, H. M., and Gollapudi, B. B. (2010 Oct) Epigenetics and chemical safety assessment. *Mutat Res*, **705**, 83–95.

[3] Phillips, J. M., Yamamoto, Y., Negishi, M., Maronpot, R. R., and Goodman, J. I. (2007 Mar) Orphan nuclear receptor constitutive active/androstane receptor-mediated alterations in dna methylation during phenobarbital promotion of liver tumorigenesis. *Toxicol Sci*, **96**, 72–82.

[4] Thomson, J. P., Hunter, J. M., Lempiäinen, H., Müller, A., Terranova, R., Moggs, J. G., and Meehan, R. R. (2013) Dynamic changes in 5-hydroxymethylation signatures underpin early and late events in drug exposed liver. *Nucleic Acids Res*, **41**, 5639–54.

[5] Cunningham, M. L. (1996 Sep) Role of increased dna replication in the carcinogenic risk of nonmutagenic chemical carcinogens. *Mutat Res*, **365**, 59–69.

[6] Williams, G. M. (2008 Aug 15) Application of mode-of-action considerations in human cancer risk assessment. *Toxicol Lett*, **180**, 75–80.

[7] Dorato, M. A. and Engelhardt, J. A. (2005) The no-observed-adverse-effect-level in drug safety evaluations: Use, issues, and definition(s). *Regulatory Toxicology and Pharmacology*, **42**, 265 – 274.

[8] Olson, H., et al. (2000) Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regul Toxicol Pharmacol*, **32**, 56–67.

[9] Holsapple, M. P., Pitot, H. C., Cohen, S. M., Cohen, S. H., Boobis, A. R., Klaunig, J. E., Pastoor, T., Dellarco, V. L., and Dragan, Y. P. (2006) Mode of action in relevance of rodent liver tumors to human cancer risk. *Toxicol Sci*, **89**, 51–6.

[10] Cohen, S. M., Robinson, D., and MacDonald, J. (2001) Alternative models for carcinogenicity testing. *Toxicological Sciences*, **64**, 14–19.

[11] Kola, I. and Landis, J. (2004) Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov*, **3**, 711–5.

[12] of Health, U. D., Services, H., Food, Administration, D., for Drug Evaluation, C., and CDER, R. (2010) Guidance for industry: Estimating the maximum safe starting dose in initial clinical trials for therapeutics in adult healthy volunteers. -.

[13] Stevens, J. L. (2006) Future of toxicology–mechanisms of toxicity and drug safety: where do we go from here? *Chem Res Toxicol*, **19**, 1393–401.

[14] Williams, G. M. and Iatropoulos, M. J. (2002) Alteration of liver cell function and proliferation: differentiation between adaptation and toxicity. *Toxicol Pathol*, **30**, 41–53.

[15] Hardisty, J. F. and Brix, A. E. (2005) Comparative hepatic toxicity: prechronic/chronic liver toxicity in rodents. *Toxicol Pathol*, **33**, 35–40.

[16] Jaeschke, H. (2002) Redox considerations in hepatic injury and inflammation. *Antioxid Redox Signal*, **4**, 699–700.

[17] Noor, F., Niklas, J., Müller-Vieira, U., and Heinzle, E. (2009) An integrated approach to improved toxicity prediction for the safety assessment during preclinical drug development using hep g2 cells. *Toxicol Appl Pharmacol*, **237**, 221–31.

[18] CONNEY, A. H., DAVISON, C., GASTEL, R., and BURNS, J. J. (1960) Adaptive increases in drug-metabolizing enzymes induced by phenobarbital and other drugs. *J Pharmacol Exp Ther*, **130**, 1–8.

[19] Breen, A. P. and Murphy, J. A. (1995) Reactions of oxyl radicals with dna. *Free Radic Biol Med*, **18**, 1033–77.

[20] Frenkel, K. (1992) Carcinogen-mediated oxidant formation and oxidative dna damage. *Pharmacol Ther*, **53**, 127–66.

[21] Williams, G. M. and Jeffrey, A. M. (2000) Oxidative dna damage: endogenous and chemically induced. *Regul Toxicol Pharmacol*, **32**, 283–92.

[22] HATHWAY, D. E. (2000) Toxic action/toxicity. *Biological Reviews*, **75**, 95–127.

[23] Bucher, J. R. and Portier, C. (2004) Human carcinogenic risk evaluation, part v: The national toxicology program vision for assessing the human carcinogenic hazard of chemicals. *Toxicological Sciences*, **82**, 363–366.

[24] Liu, Z., Kelly, R., Fang, H., Ding, D., and Tong, W. (2011) Comparative analysis of predictive models for nongenotoxic hepatocarcinogenicity using both toxicogenomics and quantitative structure-activity relationships. *Chem Res Toxicol*, **24**, 1062–70.

[25] Contrera, J. F., Jacobs, A. C., and DeGeorge, J. J. (1997) Carcinogenicity testing and the evaluation of regulatory requirements for pharmaceuticals. *Regulatory Toxicology and Pharmacology*, **25**, 130 – 145.

[26] Weisburger, J. H. and Williams, G. M. (1981) Carcinogen testing: current problems and new approaches. *Science*, **214**, 401–7.

[27] Heindryckx, F., Colle, I., and Van Vlierberghe, H. (2009) Experimental mouse models for hepatocellular carcinoma research. *Int J Exp Pathol*, **90**, 367–386.

[28] Lee, K. H. and Lee, B. M. (2007) Study of mutagenicities of phthalic acid and terephthalic acid using in vitro and in vivo genotoxicity tests. *Journal of Toxicology and Environmental Health, Part A*, **70**, 1329–1335, pMID: 17654251.

[29] de Serres, F. J. and Shelby, M. D. (1979) The salmonella mutagenicity assay: recommendations. *Science*, **203**, 563–5.

[30] Fielden, M. R., Brennan, R., and Gollub, J. (2007) A gene expression biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals. *Toxicol Sci*, **99**, 90–100.

[31] Moore, M. C., Coate, K. C., Winnick, J. J., An, Z., and Cherrington, A. D. (2012) Regulation of hepatic glucose uptake and storage in vivo. *Adv Nutr*, **3**, 286–94.

[32] Kmiec, Z. (2001) Cooperation of liver cells in health and disease. *Adv Anat Embryol Cell Biol*, **161**, III–XIII, 1–151.

[33] Dardevet, D., Moore, M. C., Remond, D., Everett-Grueter, C. A., and Cherrington, A. D. (2006) Regulation of hepatic metabolism by enteral delivery of nutrients. *Nutr Res Rev*, **19**, 161–73.

[34] Michalopoulos, G. K. and DeFrances, M. C. (1997) Liver regeneration. *Science*, **276**, 60–6.

[35] Oinonen, T. and Lindros, K. O. (1998) Zonation of hepatic cytochrome p-450 expression and regulation. *Biochem J*, **329 (Pt 1)**, 17–35.

[36] Thurman, R. G., Kauffman, F., and Jungermann, K. (1986) *Regulation of Hepatic Metabolism: Intra- And Intercellular Compartmentation*. Plenum Press.

[37] SEARLE, J., HARMON, B. V., BISHOP, C. J., and KERR, J. F. R. (1987) The significance of cell death by apoptosis in hepatobiliary disease. *Journal of Gastroenterology and Hepatology*, **2**, 77–96.

[38] Kawasaki, S., Makuuchi, M., Ishizone, S., Matsunami, H., Terada, M., and Kawarazaki, H. (1992) Liver regeneration in recipients and donors after transplantation. *Lancet*, **339**, 580–581.

[39] Ledda-Columbano, G. and Columbano, A. (2011) Hepatocyte growth, proliferation and experimental carcinogenesis. Monga, S. P. S. (ed.), *Molecular Pathology of Liver Diseases*, vol. 5 of *Molecular Pathology Library*, chap. 54, pp. 791–813, Springer US.

[40] Al Kholaifi, A., Amer, A., Jeffery, B., Gray, T. J. B., Roberts, R. A., and Bell, D. R. (2008 Jul) Species-specific kinetics and zonation of hepatic dna synthesis induced by ligands of pparalpha. *Toxicol Sci*, **104**, 74–85.

[41] Columbano, A. and Shinozuka, H. (1996 Aug) Liver regeneration versus direct hyperplasia. *FASEB J*, **10**, 1118–1128.

[42] Columbano, A. and Ledda-Columbano, G. (2003) Mitogenesis by ligands of nuclear receptors: an attractive model for the study of the molecular mechanisms implicated in liver growth. *Cell Death Differ*, **10**, S19–S21.

[43] Larkins, B. A., Dilkes, B. P., Dante, R. A., Coelho, C. M., Woo, Y. M., and Liu, Y. (2001) Investigating the hows and whys of dna endoreduplication. *J Exp Bot*, **52**, 183–92.

[44] Seglen, P. O. (1997) Dna ploidy and autophagic protein degradation as determinants of hepatocellular growth and survival. *Cell Biol Toxicol*, **13**, 301–15.

[45] Toyoda, H., Bregerie, O., Vallet, A., Nalpas, B., Pivert, G., Brechot, C., and Desdouets, C. (2005) Changes to hepatocyte ploidy and binuclearity profiles during human chronic viral hepatitis. *Gut*, **54**, 297–302.

[46] Gentric, G., Desdouets, C., and Celton-Morizur, S. (2012) Hepatocytes polyploidization and cell cycle control in liver physiopathology. *Int J Hepatol*, **2012**, 282430.

[47] Brodsky, W. Y. and Uryvaeva, I. V. (1977) Cell polyploidy: its relation to tissue growth and function. *Int Rev Cytol*, **50**, 275–332.

[48] Johri, B. and D'Amato, F. (1984) *Role of Polyploidy in Reproductive Organs and Tissues*, pp. 519–566. Springer Berlin Heidelberg.

[49] Gentric, G., Celton-Morizur, S., and Desdouets, C. (2012) Polyploidy and liver proliferation. *Clin Res Hepatol Gastroenterol*, **36**, 29–34.

[50] Celton-Morizur, S., Merlen, G., Couton, D., and Desdouets, C. (2010) Polyploidy and liver proliferation: central role of insulin signaling. *Cell Cycle*, **9**, 460–6.

[51] Zhimulev, I. F., Belyaeva, E. S., Semeshin, V. F., Koryakov, D. E., Demakov, S. A., Demakova, O. V., Pokholkova, G. V., and Andreyeva, E. N. (2004) Polytene chromosomes: 70 years of genetic research. *Int Rev Cytol*, **241**, 203–75.

[52] Pandit, S. K., Westendorp, B., and de Bruin, A. (2013) Physiological significance of polyploidization in mammalian cells. *Trends Cell Biol*.

[53] Celton-Morizur, S., Merlen, G., Couton, D., Margall-Ducos, G., and Desdouets, C. (2009) The insulin/akt pathway controls a specific cell division program that leads to generation of binucleated tetraploid liver cells in rodents. *J Clin Invest*, **119**, 1880–7.

[54] Gupta, S. (2000) Hepatic polyploidy and liver growth control. *Seminars in Cancer Biology*, **10**, 161 – 171.

[55] Chen, H.-Z., et al. (2012) Canonical and atypical e2fs regulate the mammalian endocycle. *Nat Cell Biol*, **14**, 1192–202.

[56] Pandit, S. K., Westendorp, B., Nantasanti, S., van Liere, E., Tooten, P. C. J., Cornelissen, P. W. A., Toussaint, M. J. M., Lamers, W. H., and de Bruin, A. (2012 Nov) E2f8 is essential for polyploidization in mammalian cells. *Nat Cell Biol*, **14**, 1181–1191.

[57] Deschênes, J., Valet, J. P., and Marceau, N. (1981) The relationship between cell volume, ploidy. and functional activity in differentiating hepatocytes. *Cell Biophys*, **3**, 321–34.

[58] Epstein, C. J. (1967) Cell size, nuclear content, and the development of polyploidy in the mammalian liver. *Proc Natl Acad Sci U S A*, **57**, 327–34.

[59] Martin, N. C., McCullough, C. T., Bush, P. G., Sharp, L., Hall, A. C., and Harrison, D. J. (2002) Functional analysis of mouse hepatocytes differing in dna content: volume, receptor expression, and effect of ifngamma. *J Cell Physiol*, **191**, 138–44.

[60] Watanabe, T. and Tanaka, Y. (1982) Age-related alterations in the size of human hepatocytes. a study of mononuclear and binucleate cells. *Virchows Arch B Cell Pathol Incl Mol Pathol*, **39**, 9–20.

[61] Marsman, D. S., Cattley, R. C., Conway, J. G., and Popp, J. A. (1988 Dec 1) Relationship of hepatic peroxisome proliferation and replicative dna synthesis to the hepatocarcinogenicity of the peroxisome proliferators di(2-ethylhexyl)phthalate and [4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio]acetic acid (wy-14,643) in rats. *Cancer Res*, **48**, 6739–6744.

[62] Melchiorri, C., Chieco, P., Zedda, A. I., Coni, P., Ledda-Columbano, G. M., and Columbano, A. (1993 Sep) Ploidy and nuclearity of rat hepatocytes after compensatory regeneration or mitogen-induced liver growth. *Carcinogenesis*, **14**, 1825–1830.

[63] Gorla, G. R., Malhi, H., and Gupta, S. (2001 Aug) Polyploidy associated with oxidative injury attenuates proliferative potential of cells. *J Cell Sci*, **114**, 2943–2951.

[64] Celton-Morizur, S. and Desdouets, C. (2010) Polyploidization of liver cells. *Adv Exp Med Biol*, **676**, 123–35.

[65] Schwarze, P. E., Pettersen, E. O., Shoaib, M. C., and Seglen, P. O. (1984) Emergence of a population of small, diploid hepatocytes during hepatocarcinogenesis. *Carcinogenesis*, **5**, 1267–1275.

[66] Sarafoff, M., Rabes, H. M., and Dörmer, P. (1986) Correlations between ploidy and initiation probability determined by dna cytophotometry in individual altered hepatic foci. *Carcinogenesis*, **7**, 1191–6.

[67] Sargent, L., Xu, Y. H., Sattler, G. L., Meisner, L., and Pitot, H. C. (1989) Ploidy and karyotype of hepatocytes isolated from enzyme-altered foci in two different protocols of multistage hepatocarcinogenesis in the rat. *Carcinogenesis*, **10**, 387–391.

[68] Duncan, A. W., Taylor, M. H., Hickey, R. D., Hanlon Newell, A. E., Lenzi, M. L., Olson, S. B., Finegold, M. J., and Grompe, M. (2010) The ploidy conveyor of mature hepatocytes as a source of genetic variation. *Nature*, **467**, 707–10.

[69] Gold, L. S., Manley, N. B., Slone, T. H., Rohrbach, L., and Garfinkel, G. B. (2005) Supplement to the carcinogenic potency database (cpdb): Results of animal bioassays published in the general literature through 1997 and by the national toxicology program in 1997–1998. *Toxicological Sciences*, **85**, 747–808.

[70] Zucman-Rossi, J., et al. (2006) Genotype-phenotype correlation in hepatocellular adenoma: new classification and relationship with hcc. *Hepatology*, **43**, 515–24.

[71] Ogawa, K. (2009) Molecular pathology of early stage chemically induced hepatocarcinogenesis. *Pathol Int*, **59**, 605–22.

[72] Chuang, S.-C., La Vecchia, C., and Boffetta, P. (2009) Liver cancer: descriptive epidemiology and risk factors other than hbv and hcv infection. *Cancer Lett*, **286**, 9–14.

[73] Hanahan, D. and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646–74.

[74] Mittal, S. and El-Serag, H. B. (2013) Epidemiology of hepatocellular carcinoma: consider the population. *J Clin Gastroenterol*, **47 Suppl**, S2–6.

[75] Sherman, M. (2010) Epidemiology of hepatocellular carcinoma. *Oncology*, **78 Suppl 1**, 7–10.

[76] Duncan, M. B. (2013) Extracellular matrix transcriptome dynamics in hepatocellular carcinoma. *Matrix Biol*.

[77] Thurman, R. E., et al. (2012) The accessible chromatin landscape of the human genome. *Nature*, **489**, 75–82.

[78] Djebali, S., et al. (2012) Landscape of transcription in human cells. *Nature*, **489**, 101–8.

[79] Wang, J., et al. (2012) Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res*, **22**, 1798–812.

[80] Arvey, A., Agius, P., Noble, W. S., and Leslie, C. (2012) Sequence and chromatin determinants of cell-type-specific transcription factor binding. *Genome Res*, **22**, 1723–34.

[81] Sanyal, A., Lajoie, B. R., Jain, G., and Dekker, J. (2012) The long-range interaction landscape of gene promoters. *Nature*, **489**, 109–13.

[82] Coulon, A., Chow, C. C., Singer, R. H., and Larson, D. R. (2013) Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nat Rev Genet*, **14**, 572–84.

[83] Vernot, B., Stergachis, A. B., Maurano, M. T., Vierstra, J., Neph, S., Thurman, R. E., Stamatoyannopoulos, J. A., and Akey, J. M. (2012) Personal and population genomics of human regulatory variation. *Genome Res*, **22**, 1689–97.

[84] Whitfield, T. W., Wang, J., Collins, P. J., Partridge, E. C., Aldred, S. F., Trinklein, N. D., Myers, R. M., and Weng, Z. (2012) Functional analysis of transcription factor binding sites in human promoters. *Genome Biol*, **13**, R50.

[85] Vlieghe, D., Sandelin, A., De Bleser, P. J., Vleminckx, K., Wasserman, W. W., van Roy, F., and Lenhard, B. (2006) A new generation of jaspar, the open-access repository for transcription factor binding site profiles. *Nucleic Acids Res*, **34**, D95–7.

[86] Portales-Casamar, E., Thongjuea, S., Kwon, A. T., Arenillas, D., Zhao, X., Valen, E., Yusuf, D., Lenhard, B., Wasserman, W. W., and Sandelin, A. (2010) Jaspar 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res*, **38**, D105–10.

[87] Matys, V., et al. (2006) Transfac and its module transcompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res*, **34**, D108–10.

[88] Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., and Luscombe, N. M. (2009) A census of human transcription factors: function, expression and evolution. *Nat Rev Genet*, **10**, 252–63.

[89] Brykczynska, U., Hisano, M., Erkek, S., Ramos, L., Oakeley, E. J., Roloff, T. C., Beisel, C., Schübeler, D., Stadler, M. B., and Peters, A. H. F. M. (2010) Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat Struct Mol Biol*, **17**, 679–87.

[90] Dong, X., et al. (2012) Modeling gene expression using chromatin features in various cellular contexts. *Genome Biol*, **13**, R53.

[91] Ziller, M. J., et al. (2013) Charting a dynamic dna methylation landscape of the human genome. *Nature*, **500**, 477–81.

[92] Stadler, M. B., et al. (2011) Dna-binding factors shape the mouse methylome at distal regulatory regions. *Nature*, **480**, 490–5.

[93] Wang, H., et al. (2012) Widespread plasticity in ctcf occupancy linked to dna methylation. *Genome Res*, **22**, 1680–8.

[94] Tilgner, H., Knowles, D. G., Johnson, R., Davis, C. A., Chakrabortty, S., Djebali, S., Curado, J., Snyder, M., Gingeras, T. R., and Guigó, R. (2012) Deep sequencing of subcellular rna fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncrnas. *Genome Res*, **22**, 1616–25.

[95] Gelfman, S., Cohen, N., Yearim, A., and Ast, G. (2013) Dna-methylation effect on cotranscriptional splicing is dependent on gc architecture of the exon-intron structure. *Genome Res*, **23**, 789–99.

[96] Gelfman, S. and Ast, G. (2013) When epigenetics meets alternative splicing: the roles of dna methylation and gc architecture. *Epigenomics*, **5**, 351–3.

[97] Feinberg, A. P., Ohlsson, R., and Henikoff, S. (2006 Jan) The epigenetic progenitor origin of human cancer. *Nat Rev Genet*, **7**, 21–33.

[98] Cleary, S. P., et al. (2013) Identification of driver genes in hepatocellular carcinoma by exome sequencing. *Hepatology*.

[99] Kan, Z., et al. (2013) Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. *Genome Res*.

[100] Fujimoto, A., et al. (2012) Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat Genet*, **44**, 760–4.

[101] Timp, W. and Feinberg, A. P. (2013) Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. *Nat Rev Cancer*, **13**, 497–510.

[102] Suva, M. L., Riggi, N., and Bernstein, B. E. (2013 Mar 29) Epigenetic reprogramming in cancer. *Science*, **339**, 1567–1570.

[103] Malz, M., Pinna, F., Schirmacher, P., and Breuhahn, K. (2012) Transcriptional regulators in hepatocarcinogenesis–key integrators of malignant transformation. *J Hepatol*, **57**, 186–95.

[104] Calvisi, D. F., Ladu, S., Gorden, A., Farina, M., Conner, E. A., Lee, J.-S., Factor, V. M., and Thorgeirsson, S. S. (2006) Ubiquitous activation of ras and jak/stat pathways in human hcc. *Gastroenterology*, **130**, 1117–28.

[105] Calvisi, D. F., et al. (2009) Forkhead box m1b is a determinant of rat susceptibility to hepatocarcinogenesis and sustains erk activity in human hcc. *Gut*, **58**, 679–87.

[106] Malz, M., et al. (2009) Overexpression of far upstream element binding proteins: a mechanism regulating proliferation and migration in liver cancer cells. *Hepatology*, **50**, 1130–9.

[107] Li, Z., Tuteja, G., Schug, J., and Kaestner, K. H. (2012 Jan 20) Foxa1 and foxa2 are essential for sexual dimorphism in liver cancer. *Cell*, **148**, 72–83.

[108] Chen, Y.-W., Jeng, Y.-M., Yeh, S.-H., and Chen, P.-J. (2002) p53 gene and wnt signaling in benign neoplasms: -catenin mutations in hepatic adenoma but not in focal nodular hyperplasia. *Hepatology*, **36**, 927–935.

[109] Takayasu, H., Motoi, T., Kanamori, Y., Kitano, Y., Nakanishi, H., Tange, T., Nakagawara, A., and Hashizume, K. (2002) Two case reports of childhood liver cell adenomas harboring [beta]-catenin abnormalities. *Human Pathology*, **33**, 852 – 855.

[110] Torbenson, M., Lee, J.-H., Choti, M., Gage, W., Abraham, S. C., Montgomery, E., Boitnott, J., and Wu, T.-T. (2002) Hepatic adenomas: analysis of sex steroid receptor status and the wnt signaling pathway. *Mod Pathol*, **15**, 189–96.

[111] de La Coste, A., et al. (1998) Somatic mutations of the beta-catenin gene are frequent in mouse and human hepatocellular carcinomas. *Proc Natl Acad Sci U S A*, **95**, 8847–51.

[112] Miyoshi, Y., Iwao, K., Nagasawa, Y., Aihara, T., Sasaki, Y., Imaoka, S., Murata, M., Shimano, T., and Nakamura, Y. (1998) Activation of the beta-catenin gene in primary hepatocellular carcinomas by somatic alterations involving exon 3. *Cancer Res*, **58**, 2524–7.

[113] Nhieu, J. T., Renard, C. A., Wei, Y., Cherqui, D., Zafrani, E. S., and Buendia, M. A. (1999) Nuclear accumulation of mutated beta-catenin in hepatocellular carcinoma is associated with increased cell proliferation. *Am J Pathol*, **155**, 703–10.

[114] Laurent-Puig, P., Legoix, P., Bluteau, O., Belghiti, J., Franco, D., Binot, F., Monges, G., Thomas, G., Bioulac-Sage, P., and Zucman-Rossi, J. (2001) Genetic alterations associated with hepatocellular carcinomas define distinct pathways of hepatocarcinogenesis. *Gastroenterology*, **120**, 1763–73.

[115] Laity, J. H., Chung, J., Dyson, H. J., and Wright, P. E. (2000) Alternative splicing of wilms' tumor suppressor protein modulates dna binding activity through isoform-specific dna-induced conformational changes. *Biochemistry*, **39**, 5341–8.

[116] Savkur, R. S., et al. (2003) Alternative splicing within the ligand binding domain of the human constitutive androstane receptor. *Mol Genet Metab*, **80**, 216–26.

[117] Pogribny, I. P., Rusyn, I., and Beland, F. A. (2008) Epigenetic aspects of genotoxic and non-genotoxic hepatocarcinogenesis: studies in rodents. *Environ Mol Mutagen*, **49**, 9–15.

[118] Watson, R. E. and Goodman, J. I. (2002) Effects of phenobarbital on dna methylation in gc-rich regions of hepatic dna from mice that exhibit different levels of susceptibility to liver tumorigenesis. *Toxicol Sci*, **68**, 51–58.

[119] Ueda, A., Hamadeh, H. K., Webb, H. K., Yamamoto, Y., Sueyoshi, T., Afshari, C. A., Lehmann, J. M., and Negishi, M. (2002 Jan) Diverse roles of the nuclear orphan receptor car in regulating hepatic genes in response to phenobarbital. *Mol Pharmacol*, **61**, 1–6.

[120] Bachman, A. N., Phillips, J. M., and Goodman, J. I. (2006) Phenobarbital induces progressive patterns of gc-rich and gene-specific altered dna methylation in the liver of tumor-prone b6c3f1 mice. *Toxicol Sci*, **91**, 393–405.

[121] Lempiäinen, H., et al. (2011) Phenobarbital mediates an epigenetic switch at the constitutive androstane receptor (car) target gene cyp2b10 in the liver of b6c3f1 mice. *PLoS One*, **6**, e18216.

[122] Thomson, J. P., et al. (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. *Genome Biol*, **13**, R93.

[123] Lempiainen, H., et al. (2013 Feb) Identification of dlk1-dio3 imprinted gene cluster noncoding rnas as novel candidate biomarkers for liver tumor promotion. *Toxicol Sci*, **131**, 375–386.

[124] Goldsworthy, T. L., Goldsworthy, S. M., Sprankle, C. S., and Butterworth, B. E. (1994 May) Expression of myc, fos and ha-ras associated with chemically induced cell proliferation in the rat liver. *Cell Prolif*, **27**, 269–278.

[125] Johnson, P. J., Krasner, N., Portmann, B., Eddleston, A. L., and Williams, R. (1978 Nov) Hepatocellular carcinoma in great britain: influence of age, sex, hbsag status, and aetiology of underlying cirrhosis. *Gut*, **19**, 1022–1026.

[126] Drinkwater, N. (1994) The interaction of genes and hormones in murine hepatocarcinogenesis. Cockburn, A. and Smith, L. (eds.), *Nongenotoxic Carcinogenesis*, vol. 10 of *Ernst Schering Research Foundation Workshop*, pp. 219–230, Springer Berlin Heidelberg.

[127] Keng, V. W., Largaespada, D. A., and Villanueva, A. (2012) Why men are at higher risk for hepatocellular carcinoma? *J Hepatol*, **57**, 453–4.

[128] Naugler, W. E., Sakurai, T., Kim, S., Maeda, S., Kim, K., Elsharkawy, A. M., and Karin, M. (2007 Jul 6) Gender disparity in liver cancer due to sex differences in myd88-dependent il-6 production. *Science*, **317**, 121–124.

[129] Ma, W.-L., Hsu, C.-L., Yeh, C.-C., Wu, M.-H., Huang, C.-K., Jeng, L.-B., Hung, Y.-C., Lin, T.-Y., Yeh, S., and Chang, C. (2012) Hepatic androgen receptor suppresses hepatocellular carcinoma metastasis through modulation of cell migration and anoikis. *Hepatology*, **56**, 176–85.

[130] Luoma, P. V., Heikkinen, J. E., and Ylöstalo, P. R. (1984) The effect of phenobarbital on serum hormone levels and their diurnal variations in pregnancy. *Pharmacology*, **28**, 211–5.

[131] Linnoila, M., Prinz, P. N., Wonsowicz, C. J., and Leppaluoto, J. (1980) Effect of moderate doses of ethanol and phenobarbital on pituitary and thyroid hormones and testosterone. *Br J Addict*, **75**, 207–12.

[132] Nansel, D. D., Aiyer, M. S., Meinzer, W. H., 2nd, and Bogdanove, E. M. (1979) Rapid direct effects of castration and androgen treatment on luteinizing hormone-releasing hormone-induced leutinizing hormone release in the phenobarbital-treated male rat: examination of the roles direct and indirect androgen feedback mechanisms might play in the physiological control of luteinizing hormone release. *Endocrinology*, **104**, 524–31.

[133] Shapiro, B. H., Pampori, N. A., Lapenson, D. P., and Waxman, D. J. (1994) Growth hormone-dependent and -independent sexually dimorphic regulation of phenobarbital-induced hepatic cytochromes p450 2b1 and 2b2. *Arch Biochem Biophys*, **312**, 234–9.

[134] Braeuning, A., Heubach, Y., Knorpp, T., Kowalik, M. A., Templin, M., Columbano, A., and Schwarz, M. (2011 Sep) Gender-specific interplay of signaling through beta-catenin and car in the regulation of xenobiotic-induced hepatocyte proliferation. *Toxicol Sci*, **123**, 113–122.

[135] Bucher, J. R., Shackelford, C. C., Haseman, J. K., Johnson, J. D., Kurtz, P. J., and Persing, R. L. (1994 Aug) Carcinogenicity studies of oxazepam in mice. *Fundam Appl Toxicol*, **23**, 280–297.

[136] Peraino, C., Fry, R. J., and Staffeldt, E. (1973 Oct) Brief communication: Enhancement of spontaneous hepatic tumorigenesis in c3h mice by dietary phenobarbital. *J Natl Cancer Inst*, **51**, 1349–1350.

[137] Thorpe, E. and Walker, A. I. (1973 Jun) The toxicology of dieldrin (heod). ii. comparative long-term oral toxicity studies in mice with dieldrin, ddt, phenobarbitone, -bhc and -bhc. *Food Cosmet Toxicol*, **11**, 433–442.

[138] Leonardi, G. C., Candido, S., Cervello, M., Nicolosi, D., Raiti, F., Travali, S., Spandidos, D. A., and Libra, M. (2012) The tumor microenvironment in hepatocellular carcinoma (review). *Int J Oncol*, **40**, 1733–47.

[139] Lujambio, A., et al. (2013) Non-cell-autonomous tumor suppression by p53. *Cell*, **153**, 449–60.

[140] Vucur, M., Roderburg, C., Bettermann, K., Tacke, F., Heikenwalder, M., Trautwein, C., and Luedde, T. (2010) Mouse models of hepatocarcinogenesis: what can we learn for the prevention of human hepatocellular carcinoma? *Oncotarget*, **1**, 373–8.

[141] Hann, B. and Balmain, A. (2001) Building 'validated' mouse models of human cancer. *Curr Opin Cell Biol*, **13**, 778–784.

[142] Rodríguez-Antona, C., Donato, M. T., Boobis, A., Edwards, R. J., Watts, P. S., Castell, J. V., and Gómez-Lechón, M.-J. (2002) Cytochrome p450 expression in human hepatocytes and hepatoma cell lines: molecular mechanisms that determine lower expression in cultured cells. *Xenobiotica*, **32**, 505–20.

[143] Sakata, T., Watanabe, A., Takei, N., Shiota, T., Nakatsukasa, H., Fujiwara, M., Kobayashi, M., and Nagashima, H. (1983) Effect of azathioprine and carbon tetrachloride on induction of hyperplastic liver nodule and hepatocellular carcinoma by diethylnitrosamine and n-2-fluorenylacetamide in rats. *Ann N Y Acad Sci*, **417**, 288–93.

[144] Kolaja, K. L. and Klaunig, J. E. (1997) Vitamin e modulation of hepatic focal lesion growth in mice. *Toxicol Appl Pharmacol*, **143**, 380–387.

[145] Qi, Y., Chen, X., Chan, C.-y., Li, D., Yuan, C., Yu, F., Lin, M. C., Yew, D. T., Kung, H.-F., and Lai, L. (2008) Two-dimensional differential gel electrophoresis/analysis of diethylnitrosamine induced rat hepatocellular carcinoma. *Int J Cancer*, **122**, 2682–2688.

[146] Rao, K. V. and Vesselinovitch, S. D. (1973) Age- and sex-associated diethylnitrosamine dealkylation activity of the mouse liver and hepatocarcinogenesis. *Cancer Res*, **33**, 1625–1627.

[147] Zheng, W., Lu, J. J., Luo, F., Zheng, Y., Feng, Y. j., Felix, J. C., Lauchlan, S. C., and Pike, M. C. (2000) Ovarian epithelial tumor growth promotion by follicle-stimulating hormone and inhibition of the effect by luteinizing hormone. *Gynecol Oncol*, **76**, 80–8.

[148] Tran, T. T., Medline, A., and Bruce, W. R. (1996) Insulin promotion of colon tumors in rats. *Cancer Epidemiol Biomarkers Prev*, **5**, 1013–5.

[149] Diamond, L., O'Brien, T. G., and Baird, W. M. (1980) Tumor promoters and the mechanism of tumor promotion. *Adv. Cancer Res*, **32**, 74.

[150] Nakae, D., Yoshiji, H., Mizumoto, Y., Horiguchi, K., Shiraiwa, K., Tamura, K., Denda, A., and Konishi, Y. (1992) High incidence of hepatocellular carcinomas induced by a choline deficient l-amino acid defined diet in rats. *Cancer Res*, **52**, 5042–5.

[151] James, S. J., Pogribny, I. P., Pogribna, M., Miller, B. J., Jernigan, S., and Melnyk, S. (2003) Mechanisms of dna damage, dna hypomethylation, and tumor progression in the folate/methyl-deficient rat model of hepatocarcinogenesis. *J Nutr*, **133**, 3740S–3747S.

[152] Ellinger-Ziegelbauer, H., Stuart, B., Wahle, B., Bomann, W., and Ahr, H. J. (2005) Comparison of the expression profiles induced by genotoxic and nongenotoxic carcinogens in rat liver. *Mutat Res*, **575**, 61–84.

[153] Lee, G. H. (2000 Mar-Apr) Paradoxical effects of phenobarbital on mouse hepatocarcinogenesis. *Toxicol Pathol*, **28**, 215–225.

[154] Peraino, C., Fry, R. J., and Staffeldt, E. (1971 Oct) Reduction and enhancement by phenobarbital of hepatocarcinogenesis induced in the rat by 2-acetylaminofluorene. *Cancer Res*, **31**, 1506–1512.

[155] Becker, F. F. (1982 Oct) Morphological classification of mouse liver tumors based on biological characteristics. *Cancer Res*, **42**, 3918–3923.

[156] Whysner, J., Ross, P. M., and Williams, G. M. (1996) Phenobarbital mechanistic data and risk assessment: enzyme induction, enhanced cell proliferation, and tumor promotion. *Pharmacol Ther*, **71**, 153–191.

[157] Kawamoto, T., Sueyoshi, T., Zelko, I., Moore, R., Washburn, K., and Negishi, M. (1999 Sep) Phenobarbital-responsive nuclear translocation of the receptor car in induction of the cyp2b gene. *Mol Cell Biol*, **19**, 6318–6322.

[158] Diwan, B. A., Lubet, R. A., Ward, J. M., Hrabie, J. A., and Rice, J. M. (1992 Oct) Tumor-promoting and hepatocarcinogenic effects of 1,4-bis[2-(3,5-dichloropyridyloxy)]benzene (tcpobop) in dba/2ncr and c57bl/6ncr mice and an apparent promoting effect on nasal cavity tumors but not on hepatocellular tumors in f344/ncr rats initiated with n-nitrosodiethylamine. *Carcinogenesis*, **13**, 1893–1901.

[159] Phillips, J. M., Burgoon, L. D., and Goodman, J. I. (2009 Aug) The constitutive active/androstane receptor facilitates unique phenobarbital-induced expression changes of genes involved in key pathways in precancerous liver and liver tumors. *Toxicol Sci*, **110**, 319–333.

[160] Ross, J., Plummer, S. M., Rode, A., Scheer, N., Bower, C. C., Vogel, O., Henderson, C. J., Wolf, C. R., and Elcombe, C. R. (2010) Human constitutive androstane receptor (car) and pregnane x receptor (pxr) support the hypertrophic but not the hyperplastic response to the murine nongenotoxic hepatocarcinogens phenobarbital and chlordane in vivo. *Toxicological Sciences*, **116**, 452–466.

[161] Aydinlik, H., Nguyen, T. D., Moennikes, O., Buchmann, A., and Schwarz, M. (2001 Nov 22) Selective pressure during tumor promotion by phenobarbital leads to clonal outgrowth of beta-catenin-mutated mouse liver tumors. *Oncogene*, **20**, 7812–7816.

[162] Hart, M. J., de los Santos, R., Albert, I. N., Rubinfeld, B., and Polakis, P. (1998 May 7) Downregulation of beta-catenin by human axin and its association with the apc tumor suppressor, beta-catenin and gsk3 beta. *Curr Biol*, **8**, 573–581.

[163] Morin, P. J., Sparks, A. B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B., and Kinzler, K. W. (1997 Mar 21) Activation of beta-catenin-tcf signaling in colon cancer by mutations in beta-catenin or apc. *Science*, **275**, 1787–1790.

[164] Sakanaka, C., Weiss, J. B., and Williams, L. T. (1998) Bridging of -catenin and glycogen synthase kinase-3 by axin and inhibition of -catenin-mediated transcription. *Proceedings of the National Academy of Sciences*, **95**, 3020–3023.

[165] Morin, P. J., Sparks, A. B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B., and Kinzler, K. W. (1997) Activation of -catenin-tcf signaling in colon cancer by mutations in -catenin or apc. *Science*, **275**, 1787–1790.

[166] Fattet, S., et al. (2009) Beta-catenin status in paediatric medulloblastomas: correlation of immunohistochemical expression with mutational status, genetic profiles, and clinical characteristics. *J Pathol*, **218**, 86–94.

[167] Huber, O., Korn, R., McLaughlin, J., Ohsugi, M., Herrmann, B. G., and Kemler, R. (1996 Sep) Nuclear localization of beta-catenin by interaction with transcription factor lef-1. *Mech Dev*, **59**, 3–10.

[168] Morin, P. J. (1999) -catenin signaling and cancer. *BioEssays*, **21**, 1021–1030.

[169] Mulholland, D. J., Dedhar, S., Coetzee, G. A., and Nelson, C. C. (2005 Dec) Interaction of nuclear receptors with the wnt/beta-catenin/tcf signaling axis: Wnt you like to know? *Endocr Rev*, **26**, 898–915.

[170] Buchmann, A., Stinchcombe, S., Korner, W., Hagenmaier, H., and Bock, K. W. (1994 Jun) Effects of 2,3,7,8-tetrachloro- and 1,2,3,4,6,7,8-heptachlorodibenzo-p-dioxin on the proliferation of preneoplastic liver cells in the rat. *Carcinogenesis*, **15**, 1143–1150.

[171] Calvisi, D. F., Ladu, S., Factor, V. M., and Thorgeirsson, S. S. (2004 Jun) Activation of beta-catenin provides proliferative and invasive advantages in c-myc/tgf-alpha hepatocarcinogenesis promoted by phenobarbital. *Carcinogenesis*, **25**, 901–908.

[172] Devereux, T. R., Anna, C. H., Foley, J. F., White, C. M., Sills, R. C., and Barrett, J. C. (1999 Aug 19) Mutation of beta-catenin is an early event in chemically induced mouse hepatocellular carcinogenesis. *Oncogene*, **18**, 4726–4733.

[173] Diwan, B. A., Rice, J. M., Ward, J. M., Ohshima, M., and Lynch, P. H. (1984) Inhibition by phenobarbital and lack of effect of amobarbital on the development of liver tumors induced by n-nitrosodiethylamine in juvenile b6c3f1 mice. *Cancer Lett*, **23**, 223–234.

[174] Lee, G. H., Ooasa, T., and Osanai, M. (1998) Mechanism of the paradoxical, inhibitory effect of phenobarbital on hepatocarcinogenesis initiated in infant b6c3f1 mice with diethylnitrosamine. *Cancer Res*, **58**, 1665–1669.

[175] Weghorst, C. M. and Klaunig, J. E. (1989) Phenobarbital promotion in diethylnitrosamine-initiated infant b6c3f1 mice: influence of gender. *Carcinogenesis*, **10**, 609–612.

[176] Klaunig, J. E., Pereira, M. A., Ruch, R. J., and Weghorst, C. M. (1988) Dose-response relationship of diethylnitrosamine-initiated tumors in neonatal balb/c mice: effect of phenobarbital promotion. *Toxicol Pathol*, **16**, 381–385.

[177] Goldsworthy, T. L. and Fransson-Steen, R. (2002) Quantitation of the cancer process in c57bl/6j, b6c3f1 and c3h/hej mice. *Toxicol Pathol*, **30**, 97–105.

[178] Imaoka, S., Osada, M., Minamiyama, Y., Yukimura, T., Toyokuni, S., Takemura, S., Hiroi, T., and Funae, Y. (2004) Role of phenobarbital-inducible cytochrome p450s as a source of active oxygen species in dna-oxidation. *Cancer Lett*, **203**, 117–125.

[179] Forman, B. M., Tzameli, I., Choi, H. S., Chen, J., Simha, D., Seol, W., Evans, R. M., and Moore, D. D. (1998) Androstane metabolites bind to and deactivate the nuclear receptor car-beta. *Nature*, **395**, 612–615.

[180] Wada, T., Gao, J., and Xie, W. (2009) Pxr and car in energy metabolism. *Trends Endocrinol Metab*, **20**, 273–9.

[181] Kodama, S., Koike, C., Negishi, M., and Yamamoto, Y. (2004) Nuclear receptors car and pxr cross talk with foxo1 to regulate genes that encode drug-metabolizing and gluconeogenic enzymes. *Mol Cell Biol*, **24**, 7931–40.

[182] Dong, B., et al. (2009) Activation of nuclear receptor car ameliorates diabetes and fatty liver disease. *Proc Natl Acad Sci U S A*, **106**, 18831–6.

[183] Yamamoto, Y., Moore, R., Flavell, R. A., Lu, B., and Negishi, M. (2010) Nuclear receptor car represses tnfalpha-induced cell death by interacting with the anti-apoptotic gadd45b. *PLoS One*, **5**, e10121.

[184] Mutoh, S., Osabe, M., Inoue, K., Moore, R., Pedersen, L., Perera, L., Rebolloso, Y., Sueyoshi, T., and Negishi, M. (2009) Dephosphorylation of threonine 38 is required for nuclear translocation and activation of human xenobiotic receptor car (nr1i3). *J Biol Chem*, **284**, 34785–92.

[185] Yoshinari, K., Kobayashi, K., Moore, R., Kawamoto, T., and Negishi, M. (2003 Jul 31) Identification of the nuclear receptor car:hsp90 complex in mouse liver and recruitment of protein phosphatase 2a in response to phenobarbital. *FEBS Lett*, **548**, 17–20.

[186] Mutoh, S., Sobhany, M., Moore, R., Perera, L., Pedersen, L., Sueyoshi, T., and Negishi, M. (2013) Phenobarbital indirectly activates the constitutive active androstane receptor (car) by inhibition of epidermal growth factor receptor signaling. *Sci Signal*, **6**, ra31.

[187] Zelko, I., Sueyoshi, T., Kawamoto, T., Moore, R., and Negishi, M. (2001 Apr) The peptide near the c terminus regulates receptor car nuclear translocation induced by xenochemicals in mouse liver. *Mol Cell Biol*, **21**, 2838–2846.

[188] Huang, W., Zhang, J., Washington, M., Liu, J., Parant, J. M., Lozano, G., and Moore, D. D. (2005 Jun) Xenobiotic stress induces hepatomegaly and liver tumors via the nuclear receptor constitutive androstane receptor. *Mol Endocrinol*, **19**, 1646–1653.

[189] Yamamoto, Y., Moore, R., Goldsworthy, T. L., Negishi, M., and Maronpot, R. R. (2004 Oct 15) The orphan nuclear receptor constitutive active/androstane receptor is essential for liver tumor promotion by phenobarbital in mice. *Cancer Res*, **64**, 7197–7200.

[190] Lehmann, J. M., McKee, D. D., Watson, M. A., Willson, T. M., Moore, J. T., and Kliewer, S. A. (1998) The human orphan nuclear receptor pxr is activated by compounds that regulate cyp3a4 gene expression and cause drug interactions. *J Clin Invest*, **102**, 1016–23.

[191] Tolson, A. H. and Wang, H. (2010) Regulation of drug-metabolizing enzymes by xenobiotic receptors: Pxr and car. *Adv Drug Deliv Rev*, **62**, 1238–49.

[192] Shizu, R., Benoki, S., Numakura, Y., Kodama, S., Miyata, M., Yamazoe, Y., and Yoshinari, K. (2013) Xenobiotic-induced hepatocyte proliferation associated with constitutive active/androstane receptor (car) or peroxisome proliferator-activated receptor (ppar) is enhanced by pregnane x receptor (pxr) activation in mice. *PLoS One*, **8**, e61802.

[193] Nejak-Bowen, K. and Monga, S. P. (2008 Apr) Wnt/beta-catenin signaling in hepatic organogenesis. *Organogenesis*, **4**, 92–99.

[194] Burke, Z. D., Reed, K. R., Phesse, T. J., Sansom, O. J., Clarke, A. R., and Tosh, D. (2009 Jun) Liver zonation occurs through a beta-catenin-dependent, c-myc-independent mechanism. *Gastroenterology*, **136**, 2316–2324.

[195] Micsenyi, A., Tan, X., Sneddon, T., Luo, J.-H., Michalopoulos, G. K., and Monga, S. P. S. (2004) Beta-catenin is temporally regulated during normal liver development. *Gastroenterology*, **126**, 1134–46.

[196] Monga, S. P. S., Monga, H. K., Tan, X., Mulé, K., Pediaditakis, P., and Michalopoulos, G. K. (2003) -catenin antisense studies in embryonic liver cultures: Role in proliferation, apoptosis, and lineage specification. *Gastroenterology*, **124**, 202–216.

[197] Sekine, S., Lan, B. Y.-A., Bedolli, M., Feng, S., and Hebrok, M. (2006 Apr) Liver-specific loss of beta-catenin blocks glutamine synthesis pathway activity and cytochrome p450 expression in mice. *Hepatology*, **43**, 817–825.

[198] Braeuning, A., Sanna, R., Huelsken, J., and Schwarz, M. (2009 May) Inducibility of drug-metabolizing enzymes by xenobiotics in mice with liver-specific knockout of ctnnb1. *Drug Metab Dispos*, **37**, 1138–1145.

[199] Chesire, D. R., Dunn, T. A., Ewing, C. M., Luo, J., and Isaacs, W. B. (2004 Apr 1) Identification of aryl hydrocarbon receptor as a putative wnt/beta-catenin pathway target gene in prostate cancer cells. *Cancer Res*, **64**, 2523–2533.

[200] Colletti, M., Cicchini, C., Conigliaro, A., Santangelo, L., Alonzi, T., Pasquini, E., Tripodi, M., and Amicone, L. (2009 Aug) Convergence of wnt signaling on the hnf4alpha-driven transcription in controlling liver zonation. *Gastroenterology*, **137**, 660–672.

[201] Hailfinger, S., Jaworski, M., Braeuning, A., Buchmann, A., and Schwarz, M. (2006 Mar) Zonal gene expression in murine liver: lessons from tumors. *Hepatology*, **43**, 407–414.

[202] Peifer, M., Berg, S., and Reynolds, A. B. (1994 Mar 11) A repeating amino acid motif shared by proteins with diverse cellular roles. *Cell*, **76**, 789–791.

[203] Behrens, J., von Kries, J. P., Kuhl, M., Bruhn, L., Wedlich, D., Grosschedl, R., and Birchmeier, W. (1996 Aug 15) Functional interaction of beta-catenin with the transcription factor lef-1. *Nature*, **382**, 638–642.

[204] Willert, K. and Nusse, R. (1998) Beta-catenin: a key mediator of wnt signaling. *Curr Opin Genet Dev*, **8**, 95–102.

[205] Zucman-Rossi, J., Benhamouche, S., Godard, C., Boyault, S., Grimber, G., Balabaud, C., Cunha, A. S., Bioulac-Sage, P., and Perret, C. (2007 Feb 1) Differential effects of inactivated axin1 and activated beta-catenin mutations in human hepatocellular carcinomas. *Oncogene*, **26**, 774–780.

[206] Braeuning, A., et al. (2010 Nov) Phenotype and growth behavior of residual beta-catenin-positive hepatocytes in livers of beta-catenin-deficient mice. *Histochem Cell Biol*, **134**, 469–481.

[207] Harada, N., Miyoshi, H., Murai, N., Oshima, H., Tamai, Y., Oshima, M., and Taketo, M. M. (2002 Apr 1) Lack of tumorigenesis in the mouse liver after adenovirus-mediated expression of a dominant stable mutant of beta-catenin. *Cancer Res*, **62**, 1971–1977.

[208] Rignall, B., Braeuning, A., Buchmann, A., and Schwarz, M. (2011 Jan) Tumor formation in liver of conditional beta-catenin-deficient mice exposed to a diethylnitrosamine/phenobarbital tumor promotion regimen. *Carcinogenesis*, **32**, 52–57.

[209] Zhang, X.-F., Tan, X., Zeng, G., Misse, A., Singh, S., Kim, Y., Klaunig, J. E., and Monga, S. P. S. (2010 Sep) Conditional beta-catenin loss in mice promotes chemical hepatocarcinogenesis: role of oxidative stress and platelet-derived growth factor receptor alpha/phosphoinositide 3-kinase signaling. *Hepatology*, **52**, 954–965.

[210] Talmadge, J. E., Singh, R. K., Fidler, I. J., and Raz, A. (2007) Murine models to evaluate novel and conventional therapeutic strategies for cancer. *Am J Pathol*, **170**, 793–804.

[211] Pirttiaho, H. I., Sotaniemi, E. A., Pelkonen, R. O., and Pitkanen, U. (1982) Hepatic blood flow and drug metabolism in patients on enzyme-inducing anticonvulsants. *Eur J Clin Pharmacol*, **22**, 441–445.

[212] Parzefall, W., Erber, E., Sedivy, R., and Schulte-Hermann, R. (1991) Testing for induction of dna synthesis in human hepatocyte primary cultures by rat liver tumor promoters. *Cancer Res*, **51**, 1143–1147.

[213] Hasmall, S. C. and Roberts, R. A. (1999) The perturbation of apoptosis and mitosis by drugs and xenobiotics. *Pharmacol Ther*, **82**, 63–70.

[214] IARC (2001) Phenobarbital and its sodium salt. monographs, I. (ed.), *Some thyrotropic agents*, vol. 79, pp. 161–286.

[215] Lamminpaa, A., Pukkala, E., Teppo, L., and Neuvonen, P. J. (2002) Cancer incidence among patients using antiepileptic drugs: a long-term follow-up of 28,000 patients. *Eur J Clin Pharmacol*, **58**, 137–141.

[216] La Vecchia, C. and Negri, E. (2013) A review of epidemiological data on epilepsy, phenobarbital, and risk of liver cancer. *Eur J Cancer Prev*.

[217] Ma, X., Shah, Y., Cheung, C., Guo, G. L., Feigenbaum, L., Krausz, K. W., Idle, J. R., and Gonzalez, F. J. (2007) The pregnane x receptor gene-humanized mouse: a model for investigating drug-drug interactions mediated by cytochromes p450 3a. *Drug Metab Dispos*, **35**, 194–200.

[218] Cheung, C. and Gonzalez, F. J. (2008) Humanized mouse lines and their application for prediction of human drug metabolism and toxicological risk assessment. *J Pharmacol Exp Ther*, **327**, 288–99.

[219] Gonzalez, F. J. (2007) Animal models for human risk assessment: the peroxisome proliferator-activated receptor alpha-humanized mouse. *Nutr Rev*, **65**, S2–6.

[220] Xie, W. and Evans, R. M. (2002) Pharmaceutical use of mouse models humanized for the xenobiotic receptor. *Drug Discov Today*, **7**, 509–15.

[221] Scheer, N. and Wolf, C. R. (2013) Genetically humanized mouse models of drug metabolizing enzymes and transporters and their applications. *Xenobiotica*.

[222] Scheer, N., Ross, J., Rode, A., Zevnik, B., Niehaves, S., Faust, N., and Wolf, C. R. (2008) A novel panel of mouse models to evaluate the role of human pregnane x receptor and constitutive androstane receptor in drug response. *J Clin Invest*, **118**, 3228–39.

[223] Peltz, G. (2013) Can 'humanized' mice improve drug development in the 21st century? *Trends Pharmacol Sci*, **34**, 255–60.

[224] Hamadeh, H. K., Amin, R. P., Paules, R. S., and Afshari, C. A. (2002) An overview of toxicogenomics. *Curr Issues Mol Biol*, **4**, 45–56.

[225] Ulrich, R. G. (2003) The toxicogenomics of nuclear receptor agonists. *Curr Opin Chem Biol*, **7**, 505–10.

[226] Amundson, S. A., Bittner, M., Chen, Y., Trent, J., Meltzer, P., and Fornace, A. J., Jr (1999) Fluorescent cdna microarray hybridization reveals complexity and heterogeneity of cellular genotoxic stress responses. *Oncogene*, **18**, 3666–72.

[227] Waring, J. F., Jolly, R. A., Ciurlionis, R., Lum, P. Y., Praestgaard, J. T., Morfitt, D. C., Buratto, B., Roberts, C., Schadt, E., and Ulrich, R. G. (2001) Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. *Toxicol Appl Pharmacol*, **175**, 28–42.

[228] Iida, M., Anna, C. H., Holliday, W. M., Collins, J. B., Cunningham, M. L., Sills, R. C., and Devereux, T. R. (2005) Unique patterns of gene expression changes in liver after treatment of mice for 2 weeks with different known carcinogens and non-carcinogens. *Carcinogenesis*, **26**, 689–99.

[229] Kramer, J. A., Curtiss, S. W., Kolaja, K. L., Alden, C. L., Blomme, E. A. G., Curtiss, W. C., Davila, J. C., Jackson, C. J., and Bunch, R. T. (2004) Acute molecular markers of rodent hepatic carcinogenesis identified by transcription profiling. *Chem Res Toxicol*, **17**, 463–70.

[230] Nie, A. Y., et al. (2006) Predictive toxicogenomics approaches reveal underlying molecular mechanisms of nongenotoxic carcinogenicity. *Mol Carcinog*, **45**, 914–33.

[231] Tsujimura, K., Asamoto, M., Suzuki, S., Hokaiwado, N., Ogawa, K., and Shirai, T. (2006) Prediction of carcinogenic potential by a toxicogenomic approach using rat hepatoma cells. *Cancer Sci*, **97**, 1002–10.

[232] Lord, P. G., Nie, A., and McMillian, M. (2006) Application of genomics in preclinical drug safety evaluation. *Basic Clin Pharmacol Toxicol*, **98**, 537–46.

[233] Marbach, D., Roy, S., Ay, F., Meyer, P. E., Candeias, R., Kahveci, T., Bristow, C. A., and Kellis, M. (2012) Predictive regulatory models in drosophila melanogaster by integrative inference of transcriptional networks. *Genome Res*, **22**, 1334–49.

[234] Fu, Y., Jarboe, L. R., and Dickerson, J. A. (2011) Reconstructing genome-wide regulatory network of e. coli using transcriptome data and predicted transcription factor activities. *BMC Bioinformatics*, **12**, 233.

[235] Seok, J., Kaushal, A., Davis, R. W., and Xiao, W. (2010) Knowledge-based analysis of microarrays for the discovery of transcriptional regulation relationships. *BMC Bioinformatics*, **11 Suppl 1**, S8.

[236] Bussemaker, H. J., Foat, B. C., and Ward, L. D. (2007) Predictive modeling of genome-wide mrna expression: from modules to molecules. *Annu Rev Biophys Biomol Struct*, **36**, 329–47.

[237] Stuart, J. M., Segal, E., Koller, D., and Kim, S. K. (2003) A gene-coexpression network for global discovery of conserved genetic modules. *Science*, **302**, 249–255.

[238] Butte, A. J. and Kohane, I. S. (2000) Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. *Pac Symp Biocomput*, pp. 418–29.

[239] Margolin, A. A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Dalla Favera, R., and Califano, A. (2006) Aracne: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*, **7 Suppl 1**, S7.

[240] Wang, J., Chen, B., Wang, Y., Wang, N., Garbey, M., Tran-Son-Tay, R., Berceli, S. A., and Wu, R. (2013) Reconstructing regulatory networks from the dynamic plasticity of gene expression by mutual information. *Nucleic Acids Res*, **41**, e97.

[241] Bleda, M., Medina, I., Alonso, R., De Maria, A., Salavert, F., and Dopazo, J. (2012) Inferring the regulatory network behind a gene expression experiment. *Nucleic Acids Res*, **40**, W168–72.

[242] Hecker, M., Lambeck, S., Toepfer, S., van Someren, E., and Guthke, R. (2009) Gene regulatory network inference: data integration in dynamic models-a review. *Biosystems*, **96**, 86–103.

[243] Nguyen, D. H. and D'haeseleer, P. (2006) Deciphering principles of transcription regulation in eukaryotic genomes. *Mol Syst Biol*, **2**, 2006.0012.

[244] Demicheli, R. and Coradini, D. (2011) Gene regulatory networks: a new conceptual framework to analyse breast cancer behaviour. *Ann Oncol*, **22**, 1259–65.

[245] Liao, J. C., Boscolo, R., Yang, Y.-L., Tran, L. M., Sabatti, C., and Roychowdhury, V. P. (2003) Network component analysis: reconstruction of regulatory signals in biological systems. *Proc Natl Acad Sci U S A*, **100**, 15522–15527.

[246] Gao, F., Foat, B. C., and Bussemaker, H. J. (2004) Defining transcriptional networks through integrative modeling of mrna expression and transcription factor binding data. *BMC Bioinformatics*, **5**, 31.

[247] Sanguinetti, G., Rattray, M., and Lawrence, N. D. (2006) A probabilistic dynamical model for quantitative inference of the regulatory mechanism of transcription. *Bioinformatics*, **22**, 1753–9.

[248] Hu, Z., Killion, P. J., and Iyer, V. R. (2007) Genetic reconstruction of a functional transcriptional regulatory network. *Nat Genet*, **39**, 683–687.

[249] Pinna, A., Soranzo, N., and de la Fuente, A. (2010) From knockouts to networks: establishing direct cause-effect relationships through graph analysis. *PLoS One*, **5**, e12912.

[250] Harbison, C. T., et al. (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature*, **431**, 99–104.

[251] Das, D., Nahlé, Z., and Zhang, M. Q. (2006) Adaptively inferring human transcriptional subnetworks. *Mol Syst Biol*, **2**, 2006.0029.

[252] Suzuki, H., et al. (2009 May) The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. *Nat Genet*, **41**, 553–562.

[253] Haynes, B. C., Maier, E. J., Kramer, M. H., Wang, P. I., Brown, H., and Brent, M. R. (2013) Mapping functional transcription factor networks from gene expression data. *Genome Res*.

[254] Gitter, A., Siegfried, Z., Klutstein, M., Fornes, O., Oliva, B., Simon, I., and Bar-Joseph, Z. (2009) Backup in gene regulatory networks explains differences between binding and knockout results. *Mol Syst Biol*, **5**, 276.

[255] Arnold, P., Erb, I., Pachkov, M., Molina, N., and van Nimwegen, E. (2012 Feb 15) Motevo: integrated bayesian probabilistic methods for inferring regulatory sites and motifs on multiple alignments of dna sequences. *Bioinformatics*, **28**, 487–494.

[256] Balwierz, P. J., Pachkov, M., Arnold, P., Gruber, A. J., Zavolan, M., and van Nimwegen, E. (2014) Ismara: Automated modeling of genomic signals as a democracy of regulatory motifs. *Genome Res*.

[257] Giera, S., Braeuning, A., Kohle, C., Bursch, W., Metzger, U., Buchmann, A., and Schwarz, M. (2010 May) Wnt/beta-catenin signaling activates and determines hepatic zonal expression of glutathione s-transferases in mouse liver. *Toxicol Sci*, **115**, 22–33.

[258] Odom, D. T., Dowell, R. D., Jacobsen, E. S., Nekludova, L., Rolfe, P. A., Danford, T. W., Gifford, D. K., Fraenkel, E., Bell, G. I., and Young, R. A. (2006) Core transcriptional regulatory circuitry in human hepatocytes. *Mol Syst Biol*, **2**, 2006.0017.

[259] Fearon, E. R. and Spence, J. R. (2012 Oct 9) Cancer biology: a new ring to wnt signaling. *Curr Biol*, **22**, R849–51.

[260] Conner, E. A., Lemmer, E. R., Omori, M., Wirth, P. J., Factor, V. M., and Thorgeirsson, S. S. (2000 Oct 19) Dual functions of e2f-1 in a transgenic mouse model of liver carcinogenesis. *Oncogene*, **19**, 5054–5062.

[261] Palaiologou, M., Koskinas, J., Karanikolas, M., Fatourou, E., and Tiniakos, D. G. (2012 May) E2f-1 is overexpressed and pro-apoptotic in human hepatocellular carcinoma. *Virchows Arch*, **460**, 439–446.

[262] Conner, E. A., Lemmer, E. R., Sánchez, A., Factor, V. M., and Thorgeirsson, S. S. (2003) E2f1 blocks and c-myc accelerates hepatic ploidy in transgenic mouse models. *Biochem Biophys Res Commun*, **302**, 114–20.

[263] Baena, E., Gandarillas, A., Vallespinos, M., Zanet, J., Bachs, O., Redondo, C., Fabregat, I., Martinez-A, C., and de Alboran, I. M. (2005 May 17) c-myc regulates cell size and ploidy but is not essential for postnatal proliferation in liver. *Proc Natl Acad Sci U S A*, **102**, 7286–7291.

[264] Böhm, N. and Noltemeyer, N. (1981) Excessive reversible phenobarbital induced nuclear dna-polyploidization in the growing mouse liver. *Histochemistry*, **72**, 63–74.

[265] Numoto, M., Niwa, O., Kaplan, J., Wong, K. K., Merrell, K., Kamiya, K., Yanagihara, K., and Calame, K. (1993 Aug 11) Transcriptional repressor zf5 identifies a new conserved domain in zinc finger proteins. *Nucleic Acids Res*, **21**, 3767–3775.

[266] Numoto, M., Yokoro, K., Yanagihara, K., Kamiya, K., and Niwa, O. (1995 Mar) Over-expressed zf5 gene product, a c-myc-binding protein related to gl1-kruppel protein, has a growth-suppressive activity in mouse cell lines. *Jpn J Cancer Res*, **86**, 277–283.

[267] Sobek-Klocke, I., Disque-Kochem, C., Ronsiek, M., Klocke, R., Jockusch, H., Breuning, A., Ponstingl, H., Rojas, K., Overhauser, J., and Eichenlaub-Ritter, U. (1997 Jul 15) The human gene zfp161 on 18p11.21-pter encodes a putative c-myc repressor and is homologous to murine zfp161 (chr 17) and zfp161-rs1 (x chr). *Genomics*, **43**, 156–164.

[268] El-Tanani, M., Fernig, D. G., Barraclough, R., Green, C., and Rudland, P. (2001 Nov 9) Differential modulation of transcriptional activity of estrogen receptors by direct protein-protein interactions with the t cell factor family of transcription factors. *J Biol Chem*, **276**, 41675–41682.

[269] Kouzmenko, A. P., et al. (2004 Sep 24) Wnt/beta-catenin and estrogen signaling converge in vivo. *J Biol Chem*, **279**, 40255–40258.

[270] Ignar-Trowbridge, D. M., Teng, C. T., Ross, K. A., Parker, M. G., Korach, K. S., and McLachlan, J. A. (1993 Aug) Peptide growth factors elicit estrogen receptor-dependent transcriptional activation of an estrogen-responsive element. *Mol Endocrinol*, **7**, 992–998.

[271] Levin, E. R. (2003 Mar) Bidirectional signaling between the estrogen receptor and the epidermal growth factor receptor. *Mol Endocrinol*, **17**, 309–317.

[272] Tsutsui, S., Yamamoto, R., Iishi, H., Tatsuta, M., Tsuji, M., and Terada, N. (1992) Promoting effect of ovariectomy on hepatocellular tumorigenesis induced in mice by 3'-methyl-4-dimethylaminoazobenzene. *Virchows Arch B Cell Pathol Incl Mol Pathol*, **62**, 371–375.

[273] Shimizu, I., Yasuda, M., Mizobuchi, Y., Ma, Y. R., Liu, F., Shiba, M., Horie, T., and Ito, S. (1998 Jan) Suppressive effect of oestradiol on chemical hepatocarcinogenesis in rats. *Gut*, **42**, 112–119.

[274] Yamamoto, R., Iishi, H., Tatsuta, M., Tsuji, M., and Terada, N. (1991 Aug 19) Roles of ovaries and testes in hepatocellular tumorigenesis induced in mice by 3'-methyl-4-dimethylaminoazobenzene. *Int J Cancer*, **49**, 83–88.

[275] Fontana, K., Aldrovani, M., de Paoli, F., Oliveira, H. C. F., de Campos Vidal, B., and da Cruz-Höfling, M. A. (2008) Hepatocyte nuclear phenotype: the cross-talk between anabolic androgenic steroids and exercise in transgenic mice. *Histol Histopathol*, **23**, 1367–77.

[276] Shankey, T. V., Jin, J. K., Dougherty, S., Flanigan, R. C., Graham, S., and Pyle, J. M. (1995) Dna ploidy and proliferation heterogeneity in human prostate cancers. *Cytometry*, **21**, 30–9.

[277] Bryne, J. C., Valen, E., Tang, M.-H. E., Marstrand, T., Winther, O., da Piedade, I., Krogh, A., Lenhard, B., and Sandelin, A. (2008) Jaspar, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res*, **36**, D102–6.

[278] Braeuning, A., Ittrich, C., Köhle, C., Buchmann, A., and Schwarz, M. (2007) Zonal gene expression in mouse liver resembles expression patterns of ha-ras and beta-catenin mutated hepatomas. *Drug Metab Dispos*, **35**, 503–7.

[279] Harada, N., Oshima, H., Katoh, M., Tamai, Y., Oshima, M., and Taketo, M. M. (2004) Hepatocarcinogenesis in mice with beta-catenin and ha-ras gene mutations. *Cancer Res*, **64**, 48–54.

[280] Zeller, E., Hammer, K., Kirschnick, M., and Braeuning, A. (2013) Mechanisms of ras/beta-catenin interactions. *Arch Toxicol*, **87**, 611–632.

[281] Ahmad, I., Patel, R., Liu, Y., Singh, L. B., Taketo, M. M., Wu, X.-R., Leung, H. Y., and Sansom, O. J. (2011) Ras mutation cooperates with -catenin activation to drive bladder tumourigenesis. *Cell Death Dis*, **2**, e124.

[282] Gougelet, A. and Colnot, S. (2012) A complex interplay between wnt/-catenin signalling and the cell cycle in the adult liver. *Int J Hepatol*, **2012**, 816125.

[283] Olsen, J. H., Schulgen, G., Boice, J. D., Jr, Whysner, J., Travis, L. B., Williams, G. M., Johnson, F. B., and McGee, J. O. (1995) Antiepileptic treatment and risk for hepatobiliary cancer and malignant lymphoma. *Cancer Res*, **55**, 294–7.

[284] Pirttiaho, H. I., Sotaniemi, E. A., Ahokas, J. T., and Pitkänen, U. (1978) Liver size and indices of drug metabolism in epileptics. *Br J Clin Pharmacol*, **6**, 273–8.

[285] Choi, H. S., Chung, M., Tzameli, I., Simha, D., Lee, Y. K., Seol, W., and Moore, D. D. (1997) Differential transactivation by two isoforms of the orphan nuclear hormone receptor car. *J Biol Chem*, **272**, 23565–71.

[286] Moore, L. B., Maglich, J. M., McKee, D. D., Wisely, B., Willson, T. M., Kliewer, S. A., Lambert, M. H., and Moore, J. T. (2002) Pregnane x receptor (pxr), constitutive androstane receptor (car), and benzoate x receptor (bxr) define three pharmacologically distinct classes of nuclear receptors. *Mol Endocrinol*, **16**, 977–86.

[287] Giguère, V. et al. (1999) Orphan nuclear receptors: from gene to function. *Endocrine reviews*, **20**, 689–725.

[288] Min, G., Kemper, J. K., and Kemper, B. (2002) Glucocorticoid receptor-interacting protein 1 mediates ligand-independent nuclear translocation and activation of constitutive androstane receptor in vivo. *J Biol Chem*, **277**, 26356–63.

Raphaëlle Luisier

Computational Biologist, PhD

44 Lincoln's Inn Field
WC2A 3LY
London, UK
+44 (0)744 5217 336
raphaelle.luisier@gmail.com
Swiss – 31 years old – Single

Education

Sep 10 – Sep 13 **PhD in Bioinformatics** *Basel University*, Basel, Switzerland. Graded *magna cum laude*.

Sep 07 – Jun 09 **MSc in Bioengineering & Biotechnology** *Swiss Federal Institute of Technology (EPFL), Lausanne (Switzerland).* GPA = 5.69/6.0. Ranked in top 10 students.

Sep 04 – Jun 07 **BSc in Life Sciences & Technology** *EPFL (Switzerland).* GPA = 5.19/6.0.

Research Experience

Jan 14 – Present **RESEARCH FELLOW**, *Laboratory of Computational Biology - London Research Institute - Cancer Research UK.* Mentor: Prof. Nicholas Luscombe.

Post-transcriptional regulatory networks in neurodegenerative diseases: development of bioinformatic approaches to interpretate and integrate brains-derived iCLIP and RNA-seq data in order to automatically generate new hypotheses that explain the molecular functions of proteins which bind RNA, especially 3'UTR.

Sep 10 – Sep 13 **PhD in BIOINFORMATICS**, *Computational and Systems Biology (Basel University) - Laboratory of Carcinogenesis and Epigenetics (Novartis), Switzerland.* PhD Committee: Prof. Erik Van Nimwegen (Basel University), Prof. Gerd A. Kullak-Ublick (MD, Zürich University), Dr Rémi Terranova (Novartis).

Reconstruction of gene regulatory networks underlying xenobiotic-induced nongenotoxic carcinogenesis: integration of comprehensive large-scale transcriptional and epigenetic data from *in vivo* studies and development of innovative bioinformatic approaches to interpretate large scale array-based data.

Sep 09 – Jun 10 **RESEARCH ASSISTANT**, *Laboratory of Stem Cell Bioengineering, EPFL.* Advisor: Prof. Matthias Luetolf.

Prospective isolation and characterisation of murine Mesenchymal Stem Cells (MSCs): combining microwell technology and engineered PEG-based gel technology, development of a platform intended to murine MSC prospective isolation, MSC study at single-cell level and MSC expansion. Development of software algorithms to display data sets on large-scale single cell behavior generated from time-lapse experiments

Sep 08 – Jun 09 **VISITING SCIENTIST**, Master project, *Laboratory of Tissue Repair & Regeneration, Queensland University of Technology (QUT), Australia.* Advisors: Prof. J.-A. Hubbell (EPFL), Prof. Z. Upton (QUT) & Prof. G. Pettet (QUT).

Mathematical modeling of the re-epithelialization of a 3 Dimensional Human Skin Equivalent (3D HSE): multidisciplinary project dedicated to the development of a mathematical model of the re-epithelialisation of the skin based on *in vitro* characterisation of the wound healing of a 3D HSE using ordinary differential equations.

Jun 08 – Aug 08 **MICS INTERNSHIP**, *Laboratory of Microsystem 2, EPFL.* Advisors: Prof. Edoardo Charbon & Dr. Emile Dupont.

Complete on-chip immunoassay: towards a Lab-on-chip system: design and fabrication of an hydrophilic microchannel for complete on-chip immunoassay.

Computer skills

Programming languages: Extensive knowledge of R and MATLAB; working knowledge of C/C++, Java, Perl/ Python, Shell scripting; reporting using LateX.

Bioinformatic tools	Extensive knowledge of Cytoscape and BioConductor modules related to gene expression, methylation, and sequencing data analysis; frequent usage of Ingenuity pathway analysis; working knowledge of Spotfire.
High-content data analysis	Expert in linear modeling, singular value decomposition and dimensionality reduction, clustering, interpretation and visualisation of large scale array-based data including transcriptomics (Affymetrix), miRNA (Agilent and Affymetrix) and methylation (Nimblegen); good understanding of RNA-seq and ChIP-seq data analysis.
Operating systems	Proficiency in Linux, MacOS and Windows. Experienced in high performance computing.

Laboratory skills

Flow cytometry (cell sorting and screening)

Live-cell imaging and image analysis

Primary cell cultures, cell line cultures, immunohistochemistry, *in vitro* 3 dimensional skin tissue engineering

Microfuildics

Teaching and Managerial Experiences

2013 **Supervision of bachelor-level intern** for 8 months.

2004 – 2008 (150hrs/year) **High school teacher** part time during my studies at EPFL, *Lycee College La Planta, Sion - Lycee College Les Creusets, Sion*. Areas of Teaching: biology, chemistry, biochemistry and mathematics.

Award

2009 Price of the best master project in Bioengineering and Biotechnology, *EPFL*

Publications

2014 R Luisier, EB Unterberger, JI Goodman, M Schwarz, JG Moggs, R Terranova, E van Nimwegen. Computational modeling identifies key gene regulatory interactions underlying phenobarbital-mediated tumor promotion, *Nucleic Acid Research* (2014).

2014 R Luisier, H Lempiainen, N Scherbichler, A Braeuning, M Geissler, V Dubost, et al. Phenobarbital Induces Cell Cycle Transcriptional Responses in Mouse Liver Humanized for Constitutive Androstane and Pregnane X Receptors, *Toxicological Sciences* (2014).

2014 EB Unterberger, J Eichner, C Wrzodek, H Lempiäinen, R Luisier, R Terranova, U Metzger, S Plummer, T Knorpp, A Braeuning, JG Moggs, MF Templin, V Honndorf, M Piotto, A Zell, M Schwarz. Ha-ras and β-catenin oncoproteins orchestrate metabolic programs in mouse liver tumors, *International Journal of Cancer* (2014).

2012 H Lempiäinen, R Luisier, A Müller, P Marc, D Heard, F Bolognani, P Moulin, P Couttet, O Grenet, J Marlowe, JG Moggs, R Terranova. Epigenomics – Impact for Drug Safety Sciences, *Toxicology and Epigenetics* (2012).

2012 H Lempiäinen, P Couttet, F Bolognani, A Müller, V Dubost, R Luisier, et al. Identification of Dlk1-Dio3 imprinted gene cluster non-coding RNAs as novel candidate biomarkers for liver tumor promotion, *Toxicological Science* (2012).

2010 EP Dupont, R Luisier and MAM Gijs. NOA 63 as UV-curable material for fabrication of microfluidic channels with native hydrophilicity, *Microelectronic Engineering* (2010).

Languages

French	Mother tong
English	C1 (Common European Framework of Reference for Languages)
German	A2 (Common European Framework of Reference for Languages)